

11

Db 51 GAGGACACTACACCATCGTGCCTGCTATCGCGAGTTCTACAAACAG 4

RESULT 2

US-09-328-352-1502/c
; Sequence 1502, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1502
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1502

Alignment Scores:
Pred. No.: 25 Length: 735
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x US-09-328-352-1502 (1-735)

QY 129 ValSerValLeuSerAlaValThr 136

Db 235 GTGAGTGTGTTAGCGCGGTAACA 212

RESULT 3

US-08-584-226-1/c
; Sequence 1, Application US/08584226
; Patent No. 5798240
; GENERAL INFORMATION:
; APPLICANT: Martinis, Susan A.
; APPLICANT: Sassanfar, Mandana
; APPLICANT: Kim, Sunghoon
; APPLICANT: Lee, Sang Ho
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL METHIONYL-tRNA
; TITLE OF INVENTION: SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/584,226
; FILING DATE:
; CLASSIFICATION: 536

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/305,766
; FILING DATE: 13-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CFI94-05Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1563
US-08-584-226-1

Alignment Scores:

Pred. No.: 76.5 Length: 2290
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 1 Gaps: 0

US-10-017-410-4 (1-275) x US-08-584-226-1 (1-2290)

QY 71 LeuValValValGlyIleGlySer 78

Db 309 CTCGTGTGTCGCGCATCGTAGT 286

RESULT 4

US-09-543-681A-823/c
; Sequence 823, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 823
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-823

Alignment Scores:

Pred. No.: 81.1 Length: 2427
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x US-09-543-681A-823 (1-2427)

QY 201 LeuLeuSerSerPheAsnPhePro 208

Db 1520 TTGCTTTCATCTTTTAAATTCCT 1497

RESULT 5

US-09-206-942-56
; Sequence 56, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08

Db 6384 CTCCTCGATGAATATGCACCTGC 6407

RESULT 9

US-09-029-047C-3
; Sequence 3, Application US/09029047C
; Patent No. 6632936
; GENERAL INFORMATION:
; APPLICANT: Carr, Antony M.
; TITLE OF INVENTION: Cell-Cycle Checkpoint Genes
; FILE REFERENCE: 27866/34132
; CURRENT APPLICATION NUMBER: US/09/029,047C
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: PCT/GB96/02197
; PRIOR FILING DATE: 1996-09-06
; PRIOR APPLICATION NUMBER: GB 9518220.0
; PRIOR FILING DATE: 1995-09-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
; LENGTH: 8022
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (585)..(7742)
; OTHER INFORMATION:
US-09-029-047C-3

Alignment Scores:
Pred. No.: 263 Length: 8022
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x US-09-029-047C-3 (1-8022)

Qy 83 PheThrLeuSerPheLeuGlyCln 90

Db 1835 TTTACATTATCTTCTCTGGACAA 1858

RESULT 10

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.24e+05 Length: 4403765
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.91% Indels: 0
DB: 3 Gaps: 0

US-10-017-410-4 (1-275) x US-09-103-840A-2 (1-4403765)

Qy 71 LeuValValGlyIleGlySer 78

Db 790626 CTGGTGGTTGGCGATTGGGTCG 790649

RESULT 11

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.25e+05 Length: 4411529
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 3 Gaps: 0

US-10-017-410-4 (1-275) x US-09-103-840A-1 (1-4411529)

Qy 71 LeuValValGlyIleGlySer 78

Db 788664 CTGGTGGTTGGCGATTGGGTCG 788687

RESULT 12

US-08-849-567A-101/c
; Sequence 101, Application US/08849567A
; Patent No. 6326174
; GENERAL INFORMATION:
; APPLICANT: Joyce, Gerald F.
; APPLICANT: Breaker, Ronald R.
; TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
; FILE REFERENCE: SCRI9438
; CURRENT APPLICATION NUMBER: US/08/849,567A
; CURRENT FILING DATE: 1997-08-25
; PRIOR APPLICATION NUMBER: PCT/US95/15580
; PRIOR FILING DATE: 1995-12-01
; PRIOR APPLICATION NUMBER: 08/472,194
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/349,023
; PRIOR FILING DATE: 1994-12-02
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA enzyme
US-08-849-567A-101

Mon Sep 20 11:04:18 2004

Alignment Scores: 21.5 Length: 50
Pred. No.: 7.00 Matches: 7
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.55% Gaps: 0
DB: 4

US-10-017-410-4 (1-275) x US-08-849-567A-101 (1-50)

QY 257 GlyValProTyrValSerLeu 263
DB 50 GGGTACCTTATGTTCTTCA 30

RESULT 13
US-08-484-322-9/c
; Sequence 9, Application US/08484322
; Patent No. 6110465
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF HYPERVARIABLE
; TITLE OF INVENTION: REGION 1 OF THE ENVELOPE 2 GENE OF ISOLATES
; TITLE OF INVENTION: OF HEPATITIS C VIRUS AND THE USE OF
; TITLE OF INVENTION: REAGENTS DERIVED FROM THESE HYPERVARIABLE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,322
; FILING DATE: June 7, 1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116US1
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: D3
US-08-484-322-9

Alignment Scores: 45.8 Length: 108
Pred. No.: 7.00 Matches: 7
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.55% Gaps: 0
DB: 3

US-10-017-410-4 (1-275) x US-08-484-322-9 (1-108)

QY 10 LeuGlnAlaGlySerSerGlu 16
DB 106 TTACAAGCTGGATCTTCTGAG 86

RESULT 14

US-09-107-532A-1916
; Sequence 1916, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A. Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELEPHONE: (781)893-8277

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1916:

SEQUENCE CHARACTERISTICS:

LENGTH: 186 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...186

SEQUENCE DESCRIPTION: SEQ ID NO: 1916:

US-09-107-532A-1916

Alignment Scores: 78.2 Length: 186
Pred. No.: 7.00 Matches: 7
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.55% Gaps: 0
DB: 4

US-10-017-410-4 (1-275) x US-09-107-532A-1916 (1-186)

QY 162 LeulleAlaGluLeuLysArg 168
DB 107 CTTATTCGCACTGACGCGG 127

RESULT 15

```

US-09-833-381-491
; Sequence 491, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 491
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-491

Alignment Scores:
Pred. No.: 98.4 Length: 235
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x US-09-833-381-491 (1-235)

Qy 220 LeuAlaAlaTyrLeuGlyCys 226
Db 89 CTGCTGCTATCTCGGATGC 109

```

Search completed: September 18, 2004, 09:16:19
Job time : 1733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 07:06:20 ; Search time 443 Seconds
(without alignments)
3133.291 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 275
Sequence: 1 MGAPHWWDQAGSEVDWC.....IGVPYVSLCANKKSSVKIT 275

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3327077 seqs, 2523723180 residues

Word size: 1
Total number of hits satisfying chosen parameters: 6645323

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10017410/runat_15092004_164737_4913/app.query.fasta_1.455
-DB=Published Applications NA -QEXT=fastap -SUFFIX=oli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTPMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10017410_@CGN_1_723_@runat_15092004_164737_4913
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-10-017-410-3
; Sequence 3, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Gravesel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(825)
US-10-017-410-3

Alignment Scores:

1	275	100.0	828	14	US-10-017-410-3	Sequence 3, Appli
2	192	69.8	4202	17	US-10-188-832-21	Sequence 21, Appl
3	192	69.8	4212	16	US-10-295-027-151	Sequence 151, App
4	85	30.9	822	15	US-10-182-447-6	Sequence 6, Appli
5	53	19.3	744	14	US-09-945-527-48	Sequence 48, Appl
6	42	15.3	4175	14	US-10-017-410-1	Sequence 1, Appli
7	11	4.0	1194	13	US-10-425-114-32295	Sequence 32295, A
8	11	4.0	1547	17	US-10-767-701-12261	Sequence 12261, A
9	9	3.3	529	13	US-10-027-632-73075	Sequence 73075, A
10	9	3.3	529	13	US-10-027-632-73075	Sequence 73075, A
11	9	3.3	529	16	US-10-027-632-312885	Sequence 312885, A
12	9	3.3	529	16	US-10-027-632-312885	Sequence 312885, A
13	9	3.3	907	13	US-10-027-632-120526	Sequence 120526, A
14	9	3.3	907	13	US-10-027-632-120527	Sequence 120527, A
15	9	3.3	907	16	US-10-027-632-120526	Sequence 120526, A
16	9	3.3	907	16	US-10-027-632-120527	Sequence 120527, A
17	9	3.3	48012	15	US-10-085-959-251	Sequence 251, App
18	9	3.3	48254	15	US-10-238-075-829	Sequence 829, App
19	8	2.9	293	9	US-09-796-692-5696	Sequence 5696, Ap
20	8	2.9	293	15	US-10-040-862-5696	Sequence 5696, Ap
21	8	2.9	293	16	US-10-057-475B-5696	Sequence 5696, Ap
22	8	2.9	293	16	US-10-154-884B-5696	Sequence 5696, Ap
23	8	2.9	293	17	US-10-764-324-5696	Sequence 5696, Ap
24	8	2.9	357	13	US-10-424-599-89571	Sequence 69571, A
25	8	2.9	360	9	US-09-796-692-2951	Sequence 2951, Ap
26	8	2.9	360	15	US-10-040-862-2951	Sequence 2951, Ap
27	8	2.9	360	16	US-10-057-475B-2951	Sequence 2951, Ap
28	8	2.9	360	16	US-10-154-884B-2951	Sequence 2951, Ap
29	8	2.9	360	17	US-10-764-324-2951	Sequence 2951, Ap
30	8	2.9	385	17	US-10-437-963-21072	Sequence 21072, A
31	8	2.9	405	13	US-10-424-599-126386	Sequence 126386, A
32	8	2.9	520	16	US-10-027-632-43442	Sequence 43442, A
33	8	2.9	520	16	US-10-027-632-43442	Sequence 43442, A
34	8	2.9	553	13	US-10-424-599-75687	Sequence 75687, A
35	8	2.9	568	13	US-10-085-783A-16383	Sequence 16383, A
36	8	2.9	568	16	US-10-242-535A-16383	Sequence 16383, A
37	8	2.9	582	13	US-10-027-632-66604	Sequence 66604, A
38	8	2.9	582	13	US-10-027-632-66605	Sequence 66605, A
39	8	2.9	582	13	US-10-027-632-66606	Sequence 66606, A
40	8	2.9	582	13	US-10-027-632-296031	Sequence 296031, A
41	8	2.9	582	13	US-10-027-632-296032	Sequence 296032, A
42	8	2.9	582	13	US-10-027-632-296033	Sequence 296033, A
43	8	2.9	582	16	US-10-027-632-66604	Sequence 66604, A
44	8	2.9	582	16	US-10-027-632-66605	Sequence 66605, A
45	8	2.9	582	16	US-10-027-632-66606	Sequence 66606, A

Pred. No.:	1.32e-281	Length:	828
Score:	275.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-017-410-4 (1-275) x US-10-017-410-3 (1-828)			
QY	1	MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys	20
DB	1	ATGGGCGCCCGCACCTGGTGGGACAGCTGCAGGCTGTAGTCGGAGGTGCAGCTGGTGC	60
QY	21	GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal	40
DB	61	GAGGACAACTACACCATCGTGCTGCTATCGCCGAGTTCTACAACAGATCAGCAATGTC	120
QY	41	LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu	60
DB	121	TTATTTTTCATTTTACCGCCCATCTGCATGTGTTGTTGATGATGATGCAACATGCTTG	180
QY	61	AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValValGlyIleGlySerValTyr	80
DB	181	AACAGTGACATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGAAATGGATCCGCTAC	240
QY	81	PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal	100
DB	241	TTCCATTTTACCTTAGTTTCTGGGTGAGATGCTTGATGAACTTGCAGTCCCTTTGGGTT	300
QY	101	LeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsn	120
DB	301	CTGATGTGTGCTTTGGCCATGTGGTTCCTCCAGAGAGGTATCTACCAAGATCTTTCGGAAT	360
QY	121	AspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAla	140
DB	361	GACAGGGTAGTGTTCAGGTGGTGGTCAAGTCTGCTGCTGCGGTTCACGAGTCCCTGGCA	420
QY	141	PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla	160
DB	421	TTTGTCAAGCTCGCATCAACAACTCTCTGATGACCCCTGGGAGTTCCTTGCACTGCA	480
QY	161	LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe	180
DB	481	CTGCTCATCGCAGAGCTAAAGAGTGTGACAACTGCTGTGTTTAACTGGGCTCTTC	540
QY	181	SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu	200
DB	541	TCGGGCTCTGTGTGGACCCCTGGCCCTGTTCTGTGGATCAGTGACCGAGCTTCTCGGAG	600
QY	201	LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu	220
DB	601	CTGCTGTATCCITTCACCTTCCCTACCTGCATGTCATGTGGCACAATCCTCATCTGCCTT	660
QY	221	AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu	240
DB	661	GCTGCTACCTGGGCTGTGTATGCTTTGGCTACTTTGATGCTGCTTCAGAGATTCCTGAG	720
QY	241	GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyr	260
DB	721	CAAGGCGCTGTCTCAAGTCTTGCCCAATGAGAAATGGCCCTTCATTTGGTGTCCCTAT	780
QY	261	ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr	275
DB	781	GTGTCCCTCTGTGTGCCAACAGAAATCATCAGTCAAGATCACG	825

RESULT 2

US-10-188-832-21

; Sequence 21, Application US/10188832

; Publication No. US20040076955A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

APPLICANT: Aziz, Natasha

APPLICANT: Eos Biotechnology, Inc.

```

; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
;
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
;
; TITLE OF INVENTION: Cancer
;
; FILE REFERENCE: 018501-002330US
;
; CURRENT APPLICATION NUMBER: US/10/188,832
;
; CURRENT FILING DATE: 2002-11-22
;
; PRIOR APPLICATION NUMBER: US 60/302,814
;
; PRIOR FILING DATE: 2001-07-03
;
; PRIOR APPLICATION NUMBER: US 60/310,099
;
; PRIOR FILING DATE: 2001-08-03
;
; PRIOR APPLICATION NUMBER: US 60/343,705
;
; PRIOR FILING DATE: 2001-11-08
;
; PRIOR APPLICATION NUMBER: US 60/350,666
;
; PRIOR FILING DATE: 2001-11-13
;
; PRIOR APPLICATION NUMBER: US 60/372,245
;
; PRIOR FILING DATE: 2002-04-12
;
; NUMBER OF SEQ ID NOS: 207
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 21
;
; LENGTH: 4202
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-10-188-832-21

```

Alignment Scores:		
Pred. No.:	6.1e-193	4202
Score:	192.00	192
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	99.82%	Indels: 0
DB:	17	Gaps: 0

US-10-017-410-4 (1-275) x US-10-188-832-21 (1-4202)

Qy	84	ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys	103
Db	326	ACCCTTAGTTCTTTGGGTGAGATGCTTGATGAACCTTGAGTCTCTTTGGGTCTGATGTGT	385
Qy	104	AlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsnAspArgGly	123
Db	386	GCTTTGGCCATGTGTTCCCCAGAAAGGTATCTACCAAGATCTTTTCGGAATGACCCGGGT	445
Qy	124	ArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLys	143
Db	446	AGGTTCAAGGTGGTGGTCAGTGTCTCTGTCGGGTACGACGTGCTGGCATTTGTCAAG	505
Qy	144	ProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle	163
Db	506	CCTGCCATCAACAACATCTCTGATGACCCCTGGGAGTTCCTTGCACTGCACCTGCTCATC	565
Qy	164	AlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeu	183
Db	566	GCAGAGTAAAGAGGTGTGCAACATGCTGTGTGTTAAGCTGGGGCTCTCTCCGGCCCTC	625
Qy	184	TrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSer	203
Db	626	TGTTGGACCTTGGCCCTGTTCTGCTGATCAGTACCGAGCTTTCTGCGAGCTGCTGTCA	685
Qy	204	SerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaAlaTyr	223
Db	686	TCCTTCAACTTCCCTACCTGCATGTCATGTGGCACATCTCTCATCTGCCTTGTGCTTCAC	745
Qy	224	LeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyPro	243
Db	746	CTGGGCTGTGTATGCTTTTGCTACTTTTGTATGCTGCCCTCAGAGATTCCTGAGCAAGGCCCT	805
Qy	244	ValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu	263
Db	806	GTTCATCAAGTTCTGGCCCAATGAAGATGGCCCTTCATTGGTGTCCCTATGTGTCCCTC	865
Qy	264	LeuCysAlaAsnLysLysSerSerValLysIleThr	275
Db	866	CTGTGTGGCAACAAGAAATCATTCAGTCAAGATACG	901

RESULT 3

US-10-295-027-151
 ; Sequence 151, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 151
 ; LENGTH: 4212
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-295-027-151

Alignment Scores: 6.11e-193 Length: 4212
 Pred. No.: 192.00 Matches: 192
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 69.82% Gaps: 0
 DB: 16
 US-10-017-410-4 (1-275) x US-10-295-027-151 (1-4212)
 QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103
 Db 336 ACCCTTAGTTTCTTGGGTCAGATGCTTGAACCTTGACGTCCTTTGGGTTCTGATGTGT 395
 QY 104 AlaLeuAlaMetTrpPheProArgArgTyrLeuProTyrIlePheArgAsnAspArgGly 123
 Db 396 GCTTTGGCCATGTGGTTCCTCCAGAGGATCTACCAAGAATCTTCCGAATGACCGGGGT 455
 QY 124 ArgPheLysValValSerValLeuSerAlaValThrCysLeuAlaPheVallys 143
 Db 456 AGGTTCAAGGTGGTGGTCAGTGTCTGCTGGGTACGAGTGGCTGGCATTGTCAAG 515
 QY 144 ProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle 163

Db 516 CCTGCCATCAACAACATCTCTCTGATGACCCCTGGAGTTCCTTGACATGCATGCTATC 575
 QY 164 AlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeu 183
 Db 576 GCAGAGCTAAAGAGGTGTGACAAACATGCTGTGTGTTAAGCTGGGCTCTCTCTCGGGCCTC 635
 QY 184 TrpTrpThrLeuAlaLeuPheCysTrpIleSerArgAlaPheCysGluLeuLeuSer 203
 Db 636 TGGTGACCCCTGGCCCTGTTCTGCTGGATCAGTACCCGAGCTTTCTGGGAGTGTGTCA 695
 QY 204 SerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaIleTyr 223
 Db 696 TCCTTCAACTTCCCTACCTGTCATGTCATGTGGCAGATCCCTCATGCTTCTGCTCCATC 755
 QY 224 LeuGlyCysValCysPheAlaIleTyrPheAspAlaIleSerGluIleProGluGlnGlyPro 243
 Db 756 CTGGGCTGTGTGATGCTTGGCTCTTGTGATGCTGCTCAGAGATTCCTGAGCAAGGCCCT 815
 QY 244 ValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
 Db 816 GTCATCAAGTTCTGGCCCAATGAGAAATGGGCTTCAATTGGTGTCCCTATGTCCTC 875
 QY 264 LeuCysAlaAsnLysLysSerSerValLysIleThr 275
 Db 876 CTGTGTGCCAACACAGAAATCATCAGTCAAGATCACG 911
 RESULT 4
 US-10-182-447-6
 ; Sequence 6, Application US/10182447
 ; Publication No. US20030185814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFMANN, Kay
 ; APPLICANT: RADT, Marcus
 ; TITLE OF INVENTION: CERAMIDASE
 ; FILE REFERENCE: P68050S0
 ; CURRENT APPLICATION NUMBER: US/10/182,447
 ; CURRENT FILING DATE: 2002-07-29
 ; PRIOR APPLICATION NUMBER: PCT/EP01/00900
 ; PRIOR FILING DATE: 2001-01-27
 ; PRIOR APPLICATION NUMBER: DE 10003293.1
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: DE 10011392.3
 ; PRIOR FILING DATE: 2000-03-09
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 822
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-182-447-6
 Alignment Scores: 3.56e-80 Length: 822
 Pred. No.: 85.00 Matches: 85
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 30.91% Gaps: 0
 DB: 15
 US-10-017-410-4 (1-275) x US-10-182-447-6 (1-822)
 QY 165 GluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrp 184
 Db 490 GAGCTAAAGAGGTGTGACACATGCGTGTGTTAAGCTGGGCTCTTCTCGGGCTCTGG 549
 QY 185 TrpThrLeuAlaLeuPheCysTrpIleSerArgAlaPheCysGluLeuLeuSerSer 204
 Db 550 TGGACCTGGCCCTGTTCTGCTGGATCAGTACCCGAGCTTTCTGCGAGCTGCTGTCATCC 609
 QY 205 PheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaIleTyrLeu 224
 Db 610 TTCAACTTCCCTACCTGCTGCTGATGTCATGTCGACATCCCTCATCTGCTTGTGCTACCTG 669

QY 225 GlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGluGlnGlyProVal 244
Db 670 GGCTGTGTATGCTTTCCTACTTGTGCTGCCTCAGAGATTCCTCAGCAAGCCCTGTC 729
QY 245 IleLysPheTyrPro 249
Db 730 ATCAAAATTCGGCCC 744

RESULT 5

US-09-945-527-48
; Sequence 48, Application US/09945527
; Publication No. US20030055588A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20030055588A1 Nucleic Acid Molecules Encoding
; FILE REFERENCE: 35800/237985
; CURRENT APPLICATION NUMBER: US/09/945,527
; CURRENT FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-945-527-48

Alignment Scores:

Pred. No.:	2,82e-46	Length:	744
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	19.27%	Indels:	0
DB:	10	Gaps:	0

US-10-017-410-4 (1-275) x US-09-945-527-48 (1-744)

QY 1 MetGlyAlaProHisTyrTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db 277 ATGGCGCCCGCACTGTGGAGCAGCTGCAGGCTGTAGCTCGAGGTGGACTGTGTC 336
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 337 GAGGACAACTACACCATCGTGCCTGCTATCCCGAGTTCTACAAACAGCATCAGCAATGTC 396
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
Db 397 TTAATTTTCATTTTACCAGCCCATCTGCATGTGCTTGTTT 435

RESULT 6

US-10-017-410-1
; Sequence 1, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4175
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(859)
US-10-017-410-1

Alignment Scores:

Pred. No.:	6.57e-34	Length:	4175
Score:	42.00	Matches:	42
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	15.27%	Indels:	0
DB:	14	Gaps:	0

US-10-017-410-4 (1-275) x US-10-017-410-1 (1-4175)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
Db 68 GCTGGCAGTTCGAGGTGGATTGGTGGAGGACAACTACACTATCGTCTGCATTGCC 127
QY 32 GluPheTyrAsnThrIleSerAsnValLeuPheIleLeuProProIleCysMetCys 51
Db 128 GAGTTCCTACACAGCATCAGCAACGTCCTTTTTCATTTTACCTCCCATCTGCATGTC 187
QY 52 LeuPhe 53
Db 188 TTGTTTC 193

RESULT 7

US-10-425-114-32295
; Sequence 32295, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32295
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73286E07_FLI
US-10-425-114-32295

Alignment Scores:

Pred. No.:	0.149	Length:	1194
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.00%	Indels:	0
DB:	13	Gaps:	0

US-10-017-410-4 (1-275) x US-10-425-114-32295 (1-1194)

QY 30 IleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 165 ATCGAGAAATTCACAATACCAATCTCTTAATGTC 197

RESULT 8

US-10-767-701-12261
; Sequence 12261, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 12261
; LENGTH: 1547
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS34423_1
US-10-767-701-12261

Alignment Scores:
Pred. No.: 0.19 Length: 1547
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.00% Indels: 0
DB: 17 Gaps: 0

US-10-017-410-4 (1-275) x US-10-767-701-12261 (1-1547)

Qy 30 IleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 661 ATCGAGAAATCTACAATACCATCTCTAATGTC 693

RESULT 9

US-10-027-632-73075
; Sequence 73075, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73075
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(529)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-73075

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-73075 (1-529)
Qy 66 LeulleTrpThrLeuLeuValValVal 74
Db 395 TTAATCTGGACACTACTAGTAGTG 421

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-73075 (1-529)

Qy 66 LeulleTrpThrLeuLeuValValVal 74
Db 395 TTAATCTGGACACTACTAGTAGTG 421

RESULT 10
US-10-027-632-312885
; Sequence 312885, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312885
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(529)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-312885

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-312885 (1-529)
Qy 66 LeulleTrpThrLeuLeuValValVal 74
Db 395 TTAATCTGGACACTACTAGTAGTG 421

RESULT 11
US-10-027-632-73075
; Sequence 73075, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-73075 (1-529)
Qy 66 LeulleTrpThrLeuLeuValValVal 74
Db 395 TTAATCTGGACACTACTAGTAGTG 421

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73075
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(529)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-73075

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 16 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-73075 (1-529)

QY 66 LeulleTrpThrLeuValValVal 74
DB 395 TTAATCTGGACACTACTAGTAGTG 421

RESULT 12

US-10-027-632-312885
; Sequence 312885, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312885
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(529)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-312885

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0

DB: 16 Gaps: 0
US-10-017-410-4 (1-275) x US-10-027-632-312885 (1-529)

QY 66 LeulleTrpThrLeuValValVal 74
DB 395 TTAATCTGGACACTACTAGTAGTG 421

RESULT 13

US-10-027-632-120526
; Sequence 120526, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120526
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-120526

Alignment Scores:
Pred. No.: 15.2 Length: 907
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-120526 (1-907)

QY 34 TyrAsnThrIleSerAsnValLeuPhe 42
DB 156 TATAATCTATATCAATGTCTCTC 182

RESULT 14

US-10-027-632-120527
; Sequence 120527, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

Alignment Scores:
Pred. No.: 9.14 Length: 529
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0

Mon Sep 20 11:04:18 2004

Db 156 TATAATACTATATCCCAATGTCTCTTC 182
Search completed: September 18, 2004, 08:56:31
Job time : 449 secs

;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 120527
;; LENGTH: 907
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-120527

Alignment Scores:
Pred. No.: 15.2 Length: 907
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-120527 (1-907)

Qy 34 TyrAsnThrIleSerAsnValleuPhe 42
Db 156 TATAATACTATATCCCAATGTCTCTTC 182

RESULT 15
US-10-027-632-120526
; Sequence 120526, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120526
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-120526

Alignment Scores:
Pred. No.: 15.2 Length: 907
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 16 Gaps: 0

US-10-017-410-4 (1-275) x US-10-027-632-120526 (1-907)

Qy 34 TyrAsnThrIleSerAsnValleuPhe 42

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 06:56:50 ; Search time 2465 Seconds
(without alignments)
3331.481 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 275
Sequence: 1 MGAPHWMDQLQAGSSVDWC.....IGVPYVLLCANKKSSVKIT 275

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10017410/runat.15092004.164736.4870/app_query.fasta.1.455
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DNCALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410 @CGN 1.1 5180 @runat.15092004.164736.4870 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	212	77.1	681	29	AY401889 Homo sapi
2	192	69.8	681	29	AY401890 Pan trogl
3	149	54.2	449	13	BX646596 DAFZP781B
4	42	15.3	295	10	BE668106 156540 MA
5	42	15.3	357	13	BY168309 BY168309
6	42	15.3	398	14	CB707581 AMGNNUC:M
7	42	15.3	399	14	CB706492 AMGNNUC:M
8	42	15.3	432	14	CF169808 B0818D08-
9	42	15.3	434	14	CF169808 AMGNNUC:M
10	42	15.3	454	14	CB784796 AMGNNUC:T
11	42	15.3	498	12	B1848265 470659 MA
12	42	15.3	548	14	CF169369 B0812G07-
13	42	15.3	622	14	CB723138 UI-M-GH0-
14	42	15.3	633	10	BB660847 BB660847
15	42	15.3	868	14	CA976684 AGENCOURT
16	42	15.3	1022	13	BU511164 AGENCOURT
17	42	15.3	1173	11	AK085306 Mus muscu
18	39	14.2	365	13	BY117228 BY117228
19	38	13.8	797	14	CA463294 AGENCOURT
20	37	13.5	573	13	BQ081995 K-EST0028
21	36	13.1	678	13	BU203269 604153603
22	36	13.1	698	13	BU234223 603792086
23	36	13.1	715	13	BU337163 603514014
24	36	13.1	781	13	BU232554 603409105
25	36	13.1	803	14	CF593785 AGENCOURT
26	36	13.1	847	13	BU220481 603107516
27	33	12.0	401	10	BF554219 UI-R-CO-h
28	33	12.0	485	10	BF549345 UI-R-AO-a
29	33	12.0	565	14	CF115220 Shultzomi
30	31	11.3	273	10	AW424945 50710 MAR
31	29	10.5	480	14	CB728612 AMGNNUC:M
32	29	10.5	681	29	AY401891 Mus muscu
33	25	9.1	447	28	AZ998673 2M0285C18
34	23	8.4	557	28	AZ411158 1M0184B07
35	23	8.4	677	10	BB497465 BB497465
36	23	8.4	703	28	AZ323176 1M0044014
37	21	7.6	679	29	CC489407 CH240.322
38	19	6.9	325	14	CB298572 220014-re
39	19	6.9	427	9	AA900336 UI-R-E0-C
40	19	6.9	566	10	BE026641 db29d06.x
41	19	6.9	634	9	AL864302 AL864302
42	19	6.9	648	12	BJ060795 BJ060795
43	19	6.9	685	12	BJ062108 BJ062108
44	19	6.9	937	13	BX842808 BX842808
45	18	6.5	511	10	BE025879 db29d06.y

ALIGNMENTS

RESULT 1	AY401889	Homo sapiens HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,	681 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY401889	genomic survey sequence.				
DEFINITION	AY401889	AY401889.1 GI:39757875				
ACCESSION	AY401889	GSS.				
VERSION	AY401889.1	GSS.				
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 681)				

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

REFERENCE 14671302

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source Location/Qualifiers

1..681

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>681

/locus_tag="HCM1041"

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-207 Length: 681

Score: 212.00 Matches: 212

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 77.09% Indels: 0

DB: 29 Gaps: 0

US-10-017-410-4 (1-275) x AY401889 (1-681)

QY 64 IletyrleuiletrprrleuValValValGlyleGlySerValTyrPheHisPhe 83

Db 43 ATTCTTAATCTGGACCTCTTTGGTGTAGTGGAAATGGATCCGCTACTTCCATTTT 102

QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103

Db 103 ACCCTTAGTTCTTGGGTGAGTCTGTGTAAGTCACTGTCAGTCTCTTGGGTCTGATGTGT 162

QY 104 AlaLeuAlaMetTrpPheProArgAgtTyrLeuProLysilePheArgAsnAspArgGly 123

Db 163 GCTTTGGCCATGTGGTTCGCCAGAGGTATCTACCAAGATCTTTCGGAATGACCGGGT 222

QY 124 AtgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLys 143

Db 223 AGTTTCAAGTGTGGTTCAGTGTCTGTCTGCGGTTACGACGTGCCTGGCATTTGTCAAG 282

QY 144 ProAlaileAsnAsnileSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle 163

Db 283 CCTGCCATCAACAACATCTCTGATGACCTGGAGTTCCTTGCACTGCTGCTCATC 342

QY 164 AlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeu 183

Db 343 CGAGAGCTAAAGAGGTGTGACACATGCGTGTGTTTAAAGTGGGCTCTTCTCGGSCCTC 402

QY 184 TrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSer 203

Db 403 TGGTGGACCTGGCCCTGTTCTGCTGGATCAGTGACCGAGCTTTCGCGAGCTGTGTCA 462

QY 204 SerPheAsnPheProTyrLeuHisCysMetTrpHisileLeuIleCysLeuAlaIleTyr 223

Db 463 TCCTTCAACTTCCTTACCTGCACTGCATGTCGACATCTCATCTGCTGCTGCTGCTAC 522

QY 224 LeuGlyCysValCysPheAlaTyrPheAspAlaIleAserGluIleProGluGlnGlyPro 243

Db 523 CTGGGCTGTGTATGCTTTGGCTACTTTGATGTGCTGCTCAGATTCCTGAGCAAGCCCT 582

QY 244 VallileLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263

Db 593 GTCATCAAGTTCTGGCCCAATGAGAAATGGGCTTCAITGGTGTCCCTATGTGTCCCTC 642

QY 264 LeuCysAlaAsnLysLysSerSerValLysileThr 275

Db 643 CTGTGTGCCAACAAAGAAATCATCATGTCAGATCAGCAG 678

RESULT 2

AY401890 681 bp DNA linear GSS 12-DEC-2003

LOCUS Pan troglodytes HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DEFINITION AY401890

ACCESSION AY401890.1 GI:39757876

VERSION GSS.

KEYWORDS Pan troglodytes (chimpanzee)

SOURCE Pan troglodytes

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 681)

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

REFERENCE 14671302

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Location/Qualifiers

1..681

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..>681

/locus_tag="HCM1041"

ORIGIN

Alignment Scores:

Pred. No.: 6e-187 Length: 681

Score: 192.00 Matches: 192

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 69.82% Indels: 0

DB: 29 Gaps: 0

US-10-017-410-4 (1-275) x AY401890 (1-681)

QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103

Db 103 ACCCTTAGTTCTTGGTTCAGATGCTTGTATGAATTCCTTGGGTCTCTGATGTGT 162

QY 104 AlaLeuAlaMetTrpPheProArgTyrLeuProLysilePheArgAsnAspArgGly 123

Db 163 GCTTTGGCCATGTGGTTCGCCAGAGGTATCTACCAAGATCTTTCGGAATGACCGGGT 222

QY 124 ArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLys 143

Db 223 AGTTTCAAGTGTGGTTCAGTGTCTGTCTGCGGTTACGACGTGCCTGGCATTTGTCAAG 282

QY 144 ProAlaileAsnAsnileSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle 163

Db 283 CTGGGCTGTGTATGCTTTGGCTACTTTGATGTGCTGCTCAGATTCCTGAGCAAGCCCT 342

FEATURES source

Location/Qualifiers
 1. .295
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 4BOV"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from day 20 and day 40
 embryos."

ORIGIN

Alignment Scores:
 Pred. No.: 7,01e-33 Length: 295
 Score: 42.00 Matches: 42
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x BF668106 (1-295)

Qy 12 AlaGlySerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAla 31
 |||||
 Db 32 GCTGGCAGCTCGAGGTCGAGTGGTGGCAGGACACATACATCGTGGCCATGCGCC 91
 |||||
 Qy 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCys 51
 |||||
 Db 92 GAGTTCACACACATGACGACAGCTCTGTTTTCATCTTGGCGCCATCTGTCATGTC 151
 |||||
 Qy 52 LeuPhe 53
 |||||
 Db 152 TTGTTTC 157

RESULT 5
 BY168309
 LOCUS
 DEFINITION BY168309 RIKEN full-length enriched, bone marrow macrophage Mus
 musculus cDNA clone I830071C15 5', mRNA sequence.
 VERSION BY168309
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE AUTHORS
 Sukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 357)
 Okazaki, Y., Furuno, M., Katukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierzki, R. M., King, B. L., Konagaya, A.,
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

TITLE

JOURNAL
 MEDLINE
 PUBMED

COMMENT

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URU:http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Tissues were provided by David A. Hume (Depts. of Biochemistry
 and Microbiology/Parasitology Institute for Molecular Bioscience
 University of Queensland Brisbane, Q 4072 Australia) whose
 assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 source

Location/Qualifiers
 1. .357
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="I830071C15"
 /tissue_type="bone marrow"
 /cell_type="macrophage"
 /clone_lib="RIKEN full-length enriched, bone marrow
 macrophage"

ORIGIN

Alignment Scores:

Pred. No.: 8,59e-33 Length: 357
 Score: 42.00 Matches: 42
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x BY168309 (1-357)

Qy 12 AlaGlySerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAla 31
 |||||
 Db 90 GCTGGCAGTTCGAGGTCGAGTGGTGGCAGGACACATACATCTGTCGCTGCATTGCC 149
 |||||
 Qy 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCys 51
 |||||

Db 150 GAGTTCTACACAGCATCAGCAACGCTGTGTTTTCATTTTACCTCCCATCTGCATGTGC 209

QY 52 Leuphe 53
|||||

Db 210 TTGTTTC 215

RESULT 6
CB707581

LOCUS CB707581 398 bp mRNA linear EST 10-APR-2003

DEFINITION AMGNNUC:MRPE4-00375-C2-A mrpe4 (10380) Rattus norvegicus cDNA clone

ACCESSION CB707581

VERSION CB707581.1 GI:29764729

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 398)

AUTHORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00375 row: c column: 2.

FEATURES
Location/Qualifiers
1..398
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe4-00375-c2"
/tissue_type="placenta embryo"
/clone_lib="mrpe4 (10380)"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
placenta embryo day 17"

ORIGIN

Alignment Scores:
Pred. No.: 9.65e-33 Length: 398
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CB707581 (1-398)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
|||||

Db 77 GCTGGAGCTCGAGGTGGATTGGTCGAGGAACTACACCATCGTGCATCGCC 136
|||||

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51
|||||

Db 137 GAGTTCTACACAGCATCAGCAACGCTGTGTTTTCATTTACCGCCCATCTGCATGTGC 156
|||||

QY 52 Leuphe 53

Db 197 TTGTTTC 202

RESULT 7
CB706492

LOCUS CB706492 399 bp mRNA linear EST 10-APR-2003

DEFINITION AMGNNUC:MRPE4-00181-D6-A mrpe4 (10380) Rattus norvegicus cDNA clone

ACCESSION CB706492

VERSION CB706492.1 GI:29763640

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 399)

AUTHORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL Unpublished (2003)

COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00181 row: d column: 6.

FEATURES
Location/Qualifiers
1..399
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe4-00181-d6"
/tissue_type="placenta embryo"
/clone_lib="mrpe4 (10380)"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
placenta embryo day 17"

ORIGIN

Alignment Scores:
Pred. No.: 9.68e-33 Length: 399
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CB706492 (1-399)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
|||||

Db 76 GCTGGAGCTCGAGGTGGATTGGTCGAGGAACTACACCATCTGCTGCCATCGCC 135
|||||

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51
|||||

Db 136 GAGTTCTACACAGCATCAGCAACGCTGTGTTTTCATTTACCGCCCATCTGCATGTGC 195
|||||

QY 52 Leuphe 53

Db 196 TTGTTTC 201

RESULT 8
CF169808

LOCUS CF169808 432 bp mRNA linear EST 25-JUL-2003

DEFINITION B0818D08-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus

ACCESSION CF169808

VERSION CF169808.1 GI:33279357

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 432)

AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.

TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE 21429098

FURNISHED 11544199

COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0818 row: D column: 08
Seq primer: M13 Reverse
High quality sequence stop: 432

FEATURES source POLYA=No. Location/Qualifiers

1..432
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0818D08-5"
/db_xref="taxon:10090"
/clone="NIA:B0818D08 IMAGE:30469195"
/dev_stage="Newborn Kidney"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetic, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). In brief, double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
5'-PGACTAGTCTAGATCGGAGCGCGCCCTTTT-3' from 26 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 1.05e-32 Length: 432
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CF169808 (1-432)

Qy 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
Db 34 GCTGGCAGTTCGGAGGTGGATTGGTGGAGGACAACTACATATCGTCCTGCCATGGCC 93
Qy 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51
Db 94 GAGTTCACACACGATCAGCAGCTCTGTTTTCATTTTACCTCCCATCTGCATGTGC 153

Qy 52 LeuPhe 53
Db 154 TTGTTTC 159

RESULT 9
CB758986

LOCUS CB758986 434 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNNUC:MRPE3-00055-H2-A placenta embryo D17 (10379) Rattus norvegicus cDNA clone mrpe3-00055-h2 5', mRNA sequence.

ACCESSION CB758986.1 GI:29847377
VERSION CB758986
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 434)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program

JOURNAL COMMENT

Unpublished (2003)
Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00055 row: h column: 2.
Location/Qualifiers

FEATURES source

1..434
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe3-00055-h2"
/tissue_type="placenta embryo"
/clone_lib="placenta embryo D17 (10379)"
/note="Vector: pSPORT1; placenta embryo D17"

ORIGIN

Alignment Scores:
Pred. No.: 1.05e-32 Length: 434
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CB758986 (1-434)

Qy 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
Db 104 GCTGGCAGTTCGGAGGTGGATTGGTGGAGGACAACTACATATCGTCCTGCCATGGCC 163
Qy 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51
Db 164 GAGTTCACACACGATCAGCAGCTCTGTTTTCATTTTACCTCCCATCTGCATGTGC 223

Qy 52 LeuPhe 53
Db 224 TTGTTTC 229

RESULT 10
CB784796

LOCUS CB784796 454 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNNUC:TRYP1-00001-C2-A tryp1 (10582) Rattus norvegicus cDNA clone tryp1-00001-c2 5', mRNA sequence.

ACCESSION CB784796.1 GI:29873187
VERSION CB784796
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 454)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00001 row: c column: 2.
Location/Qualifiers

FEATURES source

1..454
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="tryp1-00001-c2"
/tissue_type="prostate"
/clone_lib="tryp1 (10582)"
/note="Vector: pYVA-4LL; Site 1: HindIII; Site 2: NotI; rat prostate"

ORIGIN

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:

Pred. No.: 1.36e-32 Length: 548
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CFI69369 (1-548)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31

DB 13 GCTGGCAGTTCGGAGTGGATTGGTGGAGGACAACTACATATCGTGCCTGCCATTGCC 72

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51

DB 73 GAGTTCTACAACAGCATCAGCAACGCTCTGTTTTCATTTTACCTCCCATCGCATGTGC 132

QY 52 LeuPhe 53

DB 133 TTGTTTC 138

RESULT 13

CB723138

LOCUS CB723138 622 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GHO-ceh-f-04-0-UI.r1 NIH_BMAP_GHO Mus musculus cDNA clone
IMAGE:6839525 5', mRNA sequence.

ACCESSION

CB723138

CB723138.1 GI:29780280

EST Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 622)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousef1.html>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5

Location/Qualifiers

1. .622

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6839525"

/tissue_type="Whole brain"

/dev_stage="1, 5, and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP GH0"

/note="Organ: Brain; Vector: pYX- Asc; Site_1: Ecor I;

FEATURES

source

Site 2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 1.56e-32 Length: 622
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CB723138 (1-622)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31

DB 109 GCTGGCAGTTCGGAGTGGATTGGTGGAGGACAACTACATATCGTGCCTGCCATTGCC 168

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys 51

DB 169 GAGTTCTACAACAGCATCAGCAACGCTCTGTTTTCATTTTACCTCCCATCGCATGTGC 228

QY 52 LeuPhe 53

DB 229 TTGTTTC 234

RESULT 14

CB660847

LOCUS CB660847 633 bp mRNA linear EST 26-OCT-2001
DEFINITION RIKEN full-length enriched, 0 day neonate kidney Mus musculus cDNA clone D630008P07 5', mRNA sequence.

ACCESSION

CB660847

VERSION BB660847.1 GI:16494626

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 633)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,

Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, X.,

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9226

Email: genome-res@sc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi.K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugihara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.

FEATURES

source

Location/Qualifiers
 1. 633
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="D63008BP07"
 /tissue_type="kidney"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 kidney"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in "Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN, Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGCGCGCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGATCTCGAGTAAATTAATCCCGCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:
 Pred. No.: 1-58e-32 Length: 633
 Score: 42.00 Matches: 42
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x BB660847 (1-633)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
 |||||
 Db 102 GCTGGCAGTTCGGAGGTGGATTGGTGCAGGACAACTACACTATCGTCCTGCATTGCC 161
 |||||
 QY 32 GluPheTyrAsnThrIleSerAnValLeuPhePheIleLeuProIleCysMetCys 51
 |||||
 Db 162 GAGTTCTACAACAGCATCAGACGCTTGTGTTTTCATTTCCTCCCATCGCATGTGC 221
 |||||

QY 52 LeuPhe 53

Db 222 TTGTTTC 227

RESULT 15
CA976684

LOCUS

DEFINITION CA976684 868 bp mRNA linear EST 06-JAN-2003
 AGENCOURT 8877974 NCI CCAP Mam2 Mus musculus cDNA clone
 IMAGE:6437893 5', mRNA sequence.

ACCESSION

CA976684

VERSION

CA976684.1 GI:27509338

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 868)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LLAM13955 row: g column: 14
 High quality sequence stop: 679.

FEATURES

Location/Qualifiers

1. 868

/organism="Mus musculus"

/mol_type="mRNA"

/strains="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:6437893"

/tissue_type="tumor, biopsy sample"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam2"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:
 Pred. No.: 2-22e-32 Length: 868
 Score: 42.00 Matches: 42
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CA976684 (1-868)

QY 12 AlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31

|||||

Db 45 GCTGGCAGTTCGGAGGTGGATTGGTGCAGGACAACTACACTATCGTCCTGCATTGCC 104

|||||

QY 32 GluPheTyrAsnThrIleSerAnValLeuPhePheIleLeuProIleCysMetCys 51

|||||

Db 105 GAGTTCTACAACAGCATCAGACGCTTGTGTTTTCATTTCCTCCCATCGCATGTGC 164

QY 52 LeuPhe 53

Db 165 TTGTTTC 170

Search completed: September 18, 2004, 08:47:17
 Job time : 2469 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 05:01:58 ; Search time 375 Seconds

(without alignments)
3115.345 Million cell updates/sec

Title: US-10-017-410-4

Perfect score: 1503

Sequence: 1 MGAPHWDLQAGSSEVDWC.....IGVPVYLLCANKSSVKIT 275

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame; p2n.model -DEV=xlp
-Q/cn2_1/USPTO_spool_p/US10017410/runat_15092004_164708_4681/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DICALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410 @CGN 1 1 708 @runat_15092004_164708_4681 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1503	100.0	828	6 ABS54149	Abs54149 Human cDN
2	1473	98.0	4202	7 ACC50961	Acc50961 Human bla
3	1404	93.4	4175	6 ABS54148	Abs54148 Mouse cDN
4	1390.5	92.5	822	4 AAH48641	Aah48641 Human cer
5	541	36.0	1215	4 AAK51566	Aak51566 Human pol
6	541	36.0	1631	4 AAK52550	Aak52550 Human pol
7	538.5	35.8	792	4 AAH48639	Aah48639 Human cer
8	432	28.7	537	4 ABL24043	Ab124043 Drosophil

9	307	20.4	4758	4 ABL24042	Ab124042 Drosophil
10	282.5	18.8	4071	4 ABL30432	Ab130432 Drosophil
11	216.5	14.4	801	4 AAH48640	Aah48640 Human cer
12	216.5	14.4	1063	6 ABZ111933	Abz111933 Human pol
13	216.5	14.4	2744	3 AA96499	Aa96499 cDNA enco
14	191	12.7	1239	3 AAC40178	Aac40178 Arabidops
15	186	12.4	35425	9 ADC87616	Adc87616 Human GPC
16	181	12.0	2058	4 AAH15072	Aah15072 Human cDN
17	181	12.0	4646	5 ABV27900	Abv27900 Human pro
18	181	12.0	4646	5 ABV28037	Abv28037 Human pro
19	181	12.0	4646	5 ABV22202	Abv22202 Human pro
20	181	12.0	4646	5 ABV22064	Abv22064 Human pro
21	161.5	10.7	639	5 AAS79600	Aas79600 DNA enco
22	159	10.6	636	6 ABL90053	Ab190053 Human pol
23	155	10.3	1148	3 AAF13013	Aaf13013 Aspergill
24	137.5	9.1	354	7 ABX74425	Abx74425 Human cDN
25	136	9.0	748	4 AAH07955	Aah07955 Human cDN
26	117	7.8	6755	2 AAV21511	Aav21511 Staphyloc
27	114	7.6	235033	2 AAV57926	Aav57926 Hereditar
28	111.5	7.4	1498	9 ADC87192	Adc87192 Human GPC
29	111.5	7.4	1602	5 AAS75197	Aas75197 DNA enco
30	111.5	7.4	8526	5 ABA18356	Abal8356 Human nar
31	110.5	7.4	951	7 ACC44479	Acc44479 Gene enco
32	110	7.3	1501	9 ADC87290	Adc87290 Human GPC
33	109	7.3	1147	7 ABZ52533	Abz52533 Aspergill
34	108.5	7.2	2614	9 ADC87326	Adc87326 Human GPC
35	107	7.1	237326	2 AAV57903	Aav57903 Hereditar
36	104.5	7.0	50335	8 AAD58280	Aad58280 Human tum
37	104.5	7.0	226475	8 AAD58279	Aad58279 Human tum
38	103.5	6.9	2028	4 AAF57106	Aaf57106 S. tubero
39	102.5	6.8	269223	4 AAF28554	Aaf28554 Genomic f
40	102	6.8	5154	5 AAS84859	Aas84859 DNA enco
41	102	6.8	81940	4 AAS05390	Aas05390 Human tit
42	102	6.8	81940	6 ABK64829	Abk64829 Human ben
43	101.5	6.8	1282	9 ADC87292	Adc87292 Human GPC
44	101	6.7	147724	6 ABK83566	Abk83566 Human cDN
45	100.5	6.7	1186	9 ADC87306	Adc87306 Human GPC

ALIGNMENTS

RESULT 1

ABS54149

ID ABS54149 standard; cDNA; 828 BP.

XX ABS54149;

AC ABS54149;

XX ABS54149;

DT 25-NOV-2002 (first entry)

XX Human cDNA encoding a liver tumour marker protein, CRG-L1.

DE Human; ss; gene; liver cancer; liver tumour; CRG-L1;

XX hepatocellular cancer; chromosome 9p.

KW Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX WPI; 2002-706409/76.
 DR P-PSDB; ABG32881.
 XX
 PT Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
 PT liver cancer, is differentially expressed in liver tumors relative to
 PT normal liver tissues.
 XX
 XX Claim 2; Page 8; 11pp; English.
 XX
 CC The invention relates to a polypeptide designated as CRG-L1, which is
 CC differentially expressed in liver tumors relative to normal expression
 CC in normal liver tissues, designated CRG-L1. Also included are the
 CC encoding polynucleotides (in the case of the human sequence, mapping to
 CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
 CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
 CC sequence in the diagnosis of hepatocellular cancer in tumor cells from a
 CC liver of a human or non-human animal. The CRG-L1 protein and
 CC polynucleotide are useful as diagnostic markers for a liver cancer in
 CC humans and non-human animals, and as a system for assessing putative
 CC therapeutic agents. The present sequence encodes human CRG-L1
 XX
 SQ Sequence 828 BP; 157 A; 216 C; 211 G; 244 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.47e-143 Length: 828
 Score: 1503.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x ABS54149 (1-828)

QY 1 MetGlyAlaProHisTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
 Db 1 ATGGCGCCCGCAGCTGGAGCCAGCTGAGCGCTGTAGCTGGAGTGGACTGGTGC 60
 QY 21 GluAspAsnTrpIleValProAlaIleAlaGluPheTrpAsnThrIleSerAsnVal 40
 Db 61 GAGGACAACTACACCATGCTGCTATCGCGAGTTCTACACACGATCAGCAATGTC 120
 QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTyAlaThrCysLeu 60
 Db 121 TTATTTTTCATTTTACC GCCCATCTGCATGTCTGTTGTGATGATGCAACATGCTTG 180
 QY 61 AsnSerAspIleTyLeuIleTrpThrLeuLeuValValGlyIleGlySerValTy 80
 Db 181 AACAGTGACATCTACTTAATCTGGACTCTTTTGGTTGAGTGGGAATGGATTCGCTAC 240
 QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
 Db 241 TTCCATTTTACCTTTAGTTTCTTGGGTGAGTGTGATGAATTCGAGTCTTTGGGTT 300
 QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyLeuProLysIlePheArgAsn 120
 Db 301 CTGATGTGTCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT 360
 QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
 Db 361 GACAGGGGTAGTTCAAGGTGTGTGTGTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
 Db 421 TTTGTCAAGCTGCTCCATCAACAACTCTCTGTATGACCTTGGAGTTCTTGTGCACTGCA 480
 QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLeuGlyLeuPhe 180
 Db 481 CTGCTCATCGCAGACTAAGAGGTGTGACAACTGCGTGTGTGTGTGTGTGTGTGTGTGT 540
 QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
 Db 541 TCGGGCTCTGGTGGACCTGGCCCTGTTCTCTGTGATCATGACCGAGCTTCTTGGCAG 600

QY 201 LeuLeuSerSerPheAenPheProTyLeuHisCysMetTrpHisIleLeuLeuCysLeu 220
 Db 601 CTGCTGTATCATCTTCAACTTCCCTACCTGCACTGATGTGGCACAATCTCATCTGCCTT 660
 QY 221 AlaIaTyLeuGlyCysValCysPheAlaTyPheAspAlaAlaSerGluLeuProGlu 240
 Db 661 GCTGCTACCTGGGCTGTGTATGCTTTGCTCTACTTGTATGCTGTGCTGTGCTGTGCTGAG 720
 QY 241 GlnGlyProValIleLysPheTrpProAsnGlnLysTrpAlaPheIleGlyValProTy 260
 Db 721 CAAGCCCTGTCACTCAAGTTCTGGCCCAATGAGAAATGGGCTTCAATGGTGTCCCTAT 780
 QY 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
 Db 781 GTGTGCTCTCTGTGTGGCAACAAGAAATCATCAGTCAAGATCAAG 825

RESULT 2
 ACC50961
 ID ACC50961 standard; cDNA; 4202 BP.
 XX
 AC ACC50961;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human bladder cancer associated cDNA sequence SEQ ID NO:21.
 XX
 KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2003003906-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 03-JUL-2002; 2002WO-US021338.
 XX
 PR 03-JUL-2001; 2001US-0302814P.
 PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Aziz N;
 XX
 DR WPI; 2003-201532/19.
 DR P-PSDB; ABR48156.
 XX
 PT Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX
 PS Claim 6; Page 235; 307pp; English.
 XX
 CC The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridises to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 SQ Sequence 4202 BP; 1045 A; 892 C; 927 G; 1338 T; 0 U; 0 Other;


```
Db 215 AACAGCGCATCTACTTAATATGACGCTCTAGTTGAGTGGGATGGATCTGTCTAC 274
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db 275 TTCATGCAACGCTGAGTTTCCTGGGTGAGATGCTTGATGAATGGCAATCTGTGGGTT 334
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgGlyTrpLeuProLysIlePheArgAsn 120
Db 335 CTGATGTGTGTTGGCCATGTGTTTCCAGGAGGATTTTACCAAGATCTTTCGGAT 394
QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 395 GACAGGGGACGTTCAAGGCAGTGTGTGTCTCTGTCTGCTGCAATTAACAACGTGTGGC 454
QY 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 455 TTTATCAAGCCCGCCATCAACAATATTTCCCTGATGATCTGGGACATCCATGCACTCG 514
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db 515 CTGCTTGTTCAGAGCTGAAGAGGTGTGACAATGCGGTGTGTTTAAGCTGGGCTCTTC 574
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db 575 TCTGGCTCTGGTGAATCTGCTCTCTCTGCTGATCAGCGACCAAGCCCTTCTGTGAG 634
QY 201 LeuLeuSerSerPheAsnPheProTrpLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db 635 CTGCTCTCTCTTTCATCTCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694
QY 221 AlaAlaTrpLeuGlyCysValCysPheAlaTrpPheAspAlaAlaSerGluIleProGlu 240
Db 695 GCTTCGTACCTGGGCTGTGTGTGCTTTCGCTTACTTGTGCTGCTGCTGCTGCTGCTG 754
QY 241 GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTrp 260
Db 755 CAAGGTCCAGTCATCAGATCTTGGCCGACGAGAAATGGGCTTTTATTGGTGTCCCTTAT 814
QY 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
Db 815 GTGTCCTCTGTGTGCTGCCACAGAGTGGCAGTCAAGATCAG 859
```

RESULT 4

AAH48641
ID AAH48641 standard; DNA; 822 BP.

AC AAH48641;

XX 21-SEP-2001 (first entry)

XX Human ceramidase K3 DNA.

XX Ceramidase; human; K3; antiproliferative; anticancer; anti-eczema;
KW antiporiasis; dermatological; ceramide; sphingosine; treatment; de;
KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..822

FT /*tag= a

FT /transl except= (pos:382..387,aa:Cys)

FT /note= "No stop codon given"

XX WO200155408-A1.

XX 02-AUG-2001.

XX 27-JAN-2001; 2001WO-EP000900.

XX 27-JAN-2000; 2000DE-01003293.

09-MAR-2000; 2000DE-01011392.
(MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.
Hofmann K, Conradt M;
WPI; 2001-483256/52.
P-PSDB; AAB86365.
New ceramidase containing specific structural motif, useful for
diagnosis, prevention and treatment of ceramidase defects, e.g.
ichthyosis, also in cosmetics.
Claim 3; Page 17; 31pp; German.

This invention describes novel human ceramidase (I) containing a specific
structural motif which has antiproliferative, anticancer, anti-eczema,
antiporiasis and dermatological activity. (I) cleave ceramide, resulting
in formation of sphingosine. (I), or the nucleic acid (II) encoding it,
are used for diagnosis, prevention or treatment of diseases associated
with ceramide defects, particularly altered cell proliferation (cancer)
or altered ceramide layers on the skin (neurodermatitis, eczema,
psoriasis), also for targeted modification of the permeability barrier by
ceramidase or its activators, e.g. for transcutaneous delivery of
substances. Specifically they are used for diagnosis of ichthyosis,
particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
This sequence encodes human ceramidase K3

SQ Sequence 822 BP; 153 A; 225 C; 209 G; 235 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.2e-132 Length: 822
Score: 1390.50 Matches: 256
Percent Similarity: 94.55% Conservative: 4
Best Local Similarity: 93.09% Mismatches: 14
Query Match: 92.51% Indels: 1
DB: 4 Gaps: 1

US-10-017-410-4 (1-275) x AAH48641 (1-822)

QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db 1 ATGGGCGCCCGCACCTGGTGGGACAGCTGGCAGCTGGCAGCTGGGAGTGGACATGGCGC 60
QY 21 GluAspAsnTrpThrIleValProAlaIleAlaGluPheTrpAsnThrIleSerAsnVal 40
Db 61 GAGGACAACTACACCATCGTCCCTGCTGCCCGAGTTCTATACATGATCAGCAATGTC 120
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTrpAlaThrCysLeu 60
Db 121 TTATTTTTCATTTTACCGCCCATCTGCATGTGTCTTGTTCGTCAGTATGCACATGCTTC 180
QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db 181 AACAGCGGCATCTACTTAATCTG---CTCTGTGTGTAGCGGGAATGGATCCGCTCAC 237
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db 238 TTCATGCAACCCCTTAGTTTCTCTGGTCCAGATGCTTGATGAATTCGAGTCTTTGGGTT 297
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTrpLeuProLysIlePheArgAsn 120
Db 298 CTGATGTGTGCTTCGGTCATGTGTTTCCCAAGAGGTATCTACCAAGATCTTTCCGAAT 357
QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 358 GACCAGGGTAGTTCAAGGTGGTGTCTGTGTTCTCTGCTGAGTATGACGCTGCTGCA 417
QY 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 418 TTTGTCAAGCCTGCATCAACAACATCTCTCTGATGACCTGGGAGTTCTTTCGCTGCA 477
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180

Db 478 CTGCTCATCAGAGCTAAAGAGGTGTGACACATGCGTGTGTTAAGCTGGGCTCTTC 537
 Qy 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
 Db 538 TCGGGCTCTGGTGACCTGGCCCTGTCTGTCTGGATCAGTGACCGAGCTTCTGCGAG 597
 Qy 201 LeuLeuSerSerPheAsnPheProTrpTrpLeuHisCysMetTrpHisIleLeuLeuCysLeu 220
 Db 598 CTGCTGTCTCATCTTCAACTTCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 657
 Qy 221 AlaAlaTrpLeuGlyCysValCysPheAlaTrpPheAspAlaAlaSerGluIleProGlu 240
 Db 658 GCTGCTTACTTGGCTGTGTATGCTTTCCTTACTTGTGCTGCTGCTGCTGCTGCTGCTG 717
 Qy 241 GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTrp 260
 Db 718 CAAGGCCCTGTCTCATCAAAATCTGGCCGACGAGAAATGGGCTTCTCATTTGGTGTCCCTAT 777
 Qy 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
 Db 778 GTGTCCCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822

RESULT 5

AAK51566
 ID AAK51566 standard; cDNA; 1215 BP.

AC AAK51566;

DT 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 111.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

PI WPI; 2001-476283/51.

DR P-PSDB; AAM78433.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful

PT in diagnosis and gene therapy.

XX Claim 1; Page 750-751; 6221pp; English.

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX Sequence 1215 BP; 213 A; 374 C; 309 G; 319 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,42e-45 Length: 1215
 Score: 541.00 Matches: 103
 Percent Similarity: 57.78% Conservative: 53
 Best Local Similarity: 38.15% Mismatches: 105
 Query Match: 35.99% Indels: 9
 DB: 4 Gaps: 2

US-10-017-410-4 (1-275) x AAK51566 (1-1215)

Qy 2 GlyAlaProHisTrpTrp-----AspGlnLeuGlnAlaGly-Se 14
 Db 22 GCGCAGCCGGCTGGTGGTGGCCACTGAGACAAGATGCTAGCATCTTCGCCATCAGAG 81
 Qy 14 rSerGluValAspTrpCysGluAspAsnTrpThrIleValProAlaIleAlaGluPheTy 34
 Db 82 CTCGAGGTGGACTGGTGTGAGACAACTTCAGTACTCGGAGCTGGTGGCGCATTTCTA 141
 Qy 34 rAsnThrIleSerAsnValLeuPheIleLeuProIleCysMetCysLeuPheAs 54
 Db 142 CAACAGCTTCCCAATATCCCTTCTCATCTTCGGGCCACTGATGATGCTCTCATGCA 201
 Qy 54 pGluTrpAlaThrCysLeuAsnSerAspIleTrpLeuIleTrpThrLeuLeuValValVa 74
 Db 202 CCCGTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCTGGGCTCTTCATCATCAT 261
 Qy 74 lGlyIleGlySerValTrpPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGl 94
 Db 262 AGGCCTGTTCTCCATGATTTCCACATGACGCTCAGCTTCTTGGGCCAGCTCTGGACGA 321
 Qy 94 uLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTrpLe 114
 Db 322 GATCGCCATCTGTGCTCTTGGCAGTGGCTATAGCATATGATGATGCCCGCTGCTATT 381
 Qy 114 uProLysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeuSerAl 134
 Db 382 CCCCTCCTTCTTGGGGGAGACAGGTCCTCATCCGCTGGTCTTCATCACCACGT 441
 Qy 134 aValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThrLe 154
 Db 442 GGTGACGACCTTCTGCTCTTCTCGGCCCCAGGTCAACGCTACGCTACGCTCAACAGCAT 501
 Qy 154 uGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgVa 174
 Db 502 TGCCCTGTCACATTTCTATCATCTGCTGCGAGGAGTACAGGAAGACAGCAATAAGGAGCT 561
 Qy 174 lPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSe 194
 Db 562 TCGGCACCTGATTGAGGTCTCCGTGTTTATGGGCTGTGTCTGACGAGCTGATCAG 621
 Qy 194 rAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTrpTrpLeuHisCysMetTr 214
 Db 622 TGACCGCTGTGCTTTCAGGCTTCTGGCAGAGGATTCATTTCTTCTATCTGCACAGCATCTG 681
 Qy 214 pHisIleLeuIleCysLeuAlaAlaTrpLeuGlyCysValCysPheAlaTrpPheAspAl 234
 Db 682 GCATGTGCTCATCAGCATCACCTTCCCTTATGGCATGGTCCACATGGCTTGTGGATGC 741
 Qy 234 alaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrpAl 254
 Db 742 CAACATGATGATGCCAGGTGAACCCCTCAAGCTCCGCTACTCGGCTCGGACAGATTGGCC 801

KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.
 XX Homo sapiens.
 XX WO200155408-A1.
 XX PD 02-AUG-2001.
 XX 27-JAN-2001; 2001WO-EP000900.
 XX 27-JAN-2000; 2000DE-01003293.
 XX 09-MAR-2000; 2000DE-01011392.
 XX (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.
 XX Hofmann K, Conradt M;
 XX WPI; 2001-483256/52.
 XX DR P-PSDB; AAB86363.
 XX New ceramidase containing specific structural motif, useful for
 PT diagnosis, prevention and treatment of ceramidase defects, e.g.
 PT ichthyosis, also in cosmetics.
 XX Claim 3; Page 16; 31pp; German.
 XX This invention describes novel human ceramidase (I) containing a specific
 CC structural motif which has antiproliferative, anticancer, anti-eczema,
 CC antipsoriasis and dermatological activity. (I) cleave ceramide, resulting
 CC in formation of sphingosine. (I), or the nucleic acid (II) encoding it,
 CC are used for diagnosis, prevention or treatment of diseases associated
 CC with ceramide defects, particularly altered cell proliferation (cancer)
 CC or altered ceramide layers on the skin (neurodermatitis, eczema,
 CC psoriasis), also for targeted modification of the permeability barrier by
 CC ceramidase or its activators, e.g. for transcutaneous delivery of
 CC substances. Specifically they are used for diagnosis of ichthyosis,
 CC particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
 CC This sequence encodes human ceramidase K1
 XX
 SQ Sequence 792 BP; 145 A; 241 C; 200 G; 206 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,49e-45 Length: 792
 Score: 538.50 Matches: 98
 Percent Similarity: 60.00% Conservative: 52
 Best Local Similarity: 39.20% Mismatches: 99
 Query Match: 35.83% Indels: 1
 DB: 4 Gaps: 1

US-10-017-410-4 (1-275) x AAB48639 (1-792)

QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
 DB 25 AGCTCCGAGTGGAGTGGTGGTGGAGCAACTTCCAGTACTCGAGGCTGGTGGCCGAGTTC 84
 QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
 DB 85 TACACACGTTCTCCAAATATCCCTTCTCACTTCGGCCCACTGATGATGCTCTGATG 144
 QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuValVal 73
 DB 145 CACCGGTATGCCAGAGGCTCCCGCTACATTTAGTTGTCTGGGTCTCTTCATGATC 204
 QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
 DB 205 ATAGGCGCTTCTCCCATATTTCCACATGACGCTCAGCTTCTGGCCAGCTGCTGGAC 264
 QY 94 GluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyr 113
 DB 265 GAGATGCCATCTGTGGTCTCTGGCCAGTGGCTATAGCATATGATGATGCTGCTAT 324
 QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeuSer 133

DB 325 TTCCTCTCTTCTTGGGGGAAACAGTCCCATCTCCGCTGGTCTTTCATCACCACT 384
 QY 134 AlaValThrCysLeuAlaPheValIysProAlaIleAsnAsnIleSerLeuMetThr 153
 DB 385 GTGGTCAGACCCCTTCTGTCTCTCTGGGGCCACGGTCAACGCTACGCCCTCAACAGC 444
 QY 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArg 173
 DB 445 ATTGCCCTGCACATCTCTACATCTGTGCCAGGAGTACAGGAGACCAAGCAATAAGGAG 504
 QY 174 ValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAlaLeuPheCysTrpIle 193
 DB 505 CTTCGGCACCTGATGAGTCTCCGTGGTTTATGGCTGTCTGCTACCAAGCTGGATC 564
 QY 194 SerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMet 213
 DB 565 AGTGACCGTCTGCTTTCAGCTTCTGGCAGAGGATTCATTTCTTATCTGCACAGCATC 624
 QY 214 TrpHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAsp 233
 DB 625 TGGCATGTGCTCATCAGCATCACCTTCCCTTATGGCATGGTCACCATGGCTTGGTGGAT 684
 QY 234 AlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrp 253
 DB 685 GCCAACTATGAGATGCCAGGTGAAACCCCTCAAAAGTCCGCTACTGGCTCGGACAGTTGG 744
 QY 254 AlaPheIleGlyValProTyrValSerLeu 263
 DB 745 CCC---GTGGGCTGCTCCCTACGTGGAATC 771
 RESULT 8
 ID ABL24043 standard; DNA; 537 BP.
 AC ABL24043;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23602.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US009231.
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 23602; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

XX ABL30432;
 XX AC
 XX DT 26-MAR-2002 (first entry)
 XX DE
 XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 42769.
 XX KW
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ds.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX XX
 XX PD 27-SEP-2001.
 XX XX
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX XX
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX XX
 XX PA (PEKE) PE CORP NY.
 XX XX
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX XX
 XX DR WPI; 2001-656860/75.
 XX XX
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PT interactions.
 XX PS
 XX PS Claim 1; SEQ ID NO 42769; 21pp + Sequence Listing; English.
 XX CC
 XX CC The invention relates to an isolated nucleic acid detection reagent
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
 XX CC useful in developmental biology and in elucidating cell signalling and
 XX CC cell-cell interactions in higher eukaryotes for the development of
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
 XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 XX CC ABB2072). The sequence data for this patent did not form part of the
 XX CC printed specification, but was obtained in electronic format directly
 XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ
 XX SQ Sequence 4071 BP; 1027 A; 938 C; 941 G; 1165 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,52e-18 Length: 4071
 Score: 282.50 Matches: 85
 Percent Similarity: 42.11% Conservative: 43
 Best Local Similarity: 27.96% Mismatches: 92
 Query Match: 18.80% Indels: 85
 DB: 4 Gaps: 12
 US-10-017-410-4 (1-275) x ABL30432 (1-4071)
 QY 6 TtpTtpAspGlnLeuGlnAlaGlySerSerGluValAspTtpCysGluAspAsnTyrThr 25
 DB 2704 TTCTGTGTCGGAGTTA-----TCCGTGAATCTCTGGTACTCGCAGCTCCAGGGTG 2754
 QY 26 IleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeu 45
 DB 2755 GCATTGCC-----TTCTGTTCATATGGGC 2781
 QY 46 ProProIleCysMetCysLeuPheAspGluTyr----- 56
 DB 2782 TTGGCCAGGTGTCGTGGCATGTCAGTCAGTCAGTGTCTCGAAGTCGAATTCACCTC 2841
 QY 57 -----AlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThr 69
 DB 2842 GTTACCATGTCGTGTCGTGGCAATGCCITTAGC-----GTCCTTCTGTGGAAA 2892
 QY 70 LeuLeuValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGly 89

DB 2893 TTACTCATGTGTCCACGCTGAA-GGGATTAC-----GGC 2927
 QY 90 GlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe 109
 DB 2928 GGGCATCTGGCACACCGGACTATATACATAGTAATCATCTGCGCCCTGAACAGGTTC 2987
 QY 110 ProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysValVal 129
 DB 2988 GATCCAAA-----ACCAAGCATATACGATTAAGGGC-----TTGAAGTCCCTACG 3038
 QY 130 -SerValLeuSerAlaValThrThrCysLeuAlaPheVal-----LysProAla 146
 DB 3039 CGACAGCGTCATGATCACCCTTTGACTGCAGCATTTTCATTCGGCGGACCAAGTCTCGTT 3098
 QY 146 eAsnAsnIleSerLeuMetThrLeuGlyValProCys-----ThrAlaLeuLeuIleAl 164
 DB 3099 TAATAAT-----TGCATTCAACAGCTTGCTAATTTA 3131
 QY 164 aGluLeuLys----- 167
 DB 3132 ACACGAAAAGCATGAGCGTTCTTTTCGAAATATTTCGAAATATACAACTATTATATAC 3191
 QY 168 -----ArgCysAspAsnMetArgValPheLysLeuGlyLeuPh 180
 DB 3192 CCATATATATCTTTGTTTCATACAGAGTTAGTCACAGAGGGTCTACCGCTGGCATCCG 3251
 QY 180 eSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGl 200
 DB 3252 ATCGACGACCGCTCGGGCTGTTCGGTGTCTGCTGGATCAATGACCGGATCTTCTGCGA 3311
 QY 200 uLeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLe 220
 DB 3312 GGCCTGGTCTCGATCAATTTTCATCTACCTGACGGCTTCTGGCATATTCATTTTAT 3371
 QY 220 uAlaIatyrLeuGlyCysValCysPheAlatyrPheAspAlaAlaSerGluIleProGl 240
 DB 3372 AGCCGCTACACGGTGTCTGGTCTTTTCGCTTCTTCTACGTGGAATCGGAGCTGCCCA 3431
 QY 240 uGlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTy 260
 DB 3432 GGGACGCGCTGCTGAAGTACTGCGCAAGAACAGAGTTCGAGTTTC---GGGATACCTT 3488
 QY 260 rValSerLeu 263
 DB 3489 CATTTCGATC 3498
 RESULT 11
 AAH48640
 ID AAH48640 standard; DNA; 801 BP.
 XX
 XX AC AAH48640;
 XX XX
 XX DT 21-SEP-2001 (first entry)
 XX DE Human ceramidase K2 DNA.
 XX KW Ceramidase; human; K2; antiproliferative; anticancer; anti-eczema;
 XX KW antipsoriasis; dermatological; ceramide; sphingosine; treatment; ds;
 XX KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
 XX KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.
 XX OS Homo sapiens.
 XX XX
 XX PN WO200155408-A1.
 XX XX
 XX PD 02-AUG-2001.
 XX XX
 XX PF 27-JAN-2001; 2001WO-EP000900.
 XX XX
 XX PR 27-JAN-2000; 2000DE-01003293.
 XX PR 09-MAR-2000; 2000DE-01011392.
 XX XX

PA (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.

XX Hofmann K, Conradt M;

XX WPI; 2001-483256/52.

DR P-PSDB; AAB86364.

XX

PT New ceramidease containing specific structural motif, useful for diagnosis, prevention and treatment of ceramidase defects, e.g. ichthyosis, also in cosmetics.

XX Claim 3; Page 16; 31pp; German.

XX This invention describes novel human ceramidase (I) containing a specific structural motif which has antiproliferative, anticancer, anti-eczema, antipruritis and dermatological activity. (I) cleave ceramide, resulting in formation of sphingosine. (I), or the nucleic acid (II) encoding it, are used for diagnosis, prevention or treatment of diseases associated with ceramide defects, particularly altered cell proliferation (cancer) or altered ceramide layers on the skin (neurodermatitis, eczema, psoriasis), also for targeted modification of the permeability barrier by ceramidase or its activators, e.g. for transcutaneous delivery of substances. Specifically they are used for diagnosis of ichthyosis, particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.

XX This sequence encodes human ceramidase K2

XX SQ Sequence 801 BP; 213 A; 157 C; 174 G; 257 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,61e-12 Length: 801
Score: 216.50 Matches: 79
Percent Similarity: 40.45% Conservative: 29
Best Local Similarity: 29.59% Mismatches: 105
Query Match: 14.40% Indels: 56
DB: 4 Gaps: 9

US-10-017-410-4 (1-275) x AAH48640 (1-801)

QY 14 SerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
DB 43 ACCTCCAGCTGGAGTGGTGGGAGGAGAACTACTCCGTCGACCTGTATACATCGCCGAGTTC 102
QY 34 TyrAsnThrIleSerAsnValLeuPheIleLeuProProIleCysMetCysLeuPhe 53
DB 103 TGGAAATACAGTGAATAC---CTGATCATGATTATACCTCCAA--TGTTCCGTCGCAATTC 157
QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrIleLeuValVal 73
DB 158 AGAGTGTAGACGGCTGGAAAGCGGTACATGCTTCT-TATTAGCACTCACATG 216
QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
DB 217 GTAGGAATGGGATCTGGTGGTCTCCACATGACTCTGAAATATGAAATGCAGCTATTGGAT 276
QY 94 GluLeuAlaValLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgArgTyr 113
DB 277 GAATCCCAATGATATAC---AGCTGTTCATATTTGTGTACTGCATGTTTGAATGTTTC 333
QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValValSerValLeuSer 133
DB 334 AAGATCAAGAATCTAGTAATACCTACCTCTGCTTTTACCTTAGTCTTATTCAGTTAATA 393
QY 134 AlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThr 153
DB 394 GTAACCAACAGTTTACCTTAAAGTAAAGAGCAATATTTCCATCAGTTCATGTAT----- 447
QY 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArg 173
DB 448 -----GGATGTG 456
QY 174 ValPheLysLeuGlyLeuPheSer----- 181
DB 457 GTCATTACATTAGTACTTCGATCTATTATATTGTTATCATGGTTTATCCATGGCTTAGA 516

QY 182 GlyLeuTyrTrpThr-----LeuAlaLeuPheCysTrpIleSerAsp 195
DB 517 GGACTGGGTATACATCATTCATGGGTATATTTTATTTGGATTTTATTTTGGATATAGAT 576
QY 196 ArgAlaPheCysGluLeuLeuSerSerPheAsn----- 206
DB 577 AACATATTTTGTGAGTCACTGAGGAACCTTCGAAAGAGAGGTACCACCTATCATAGGTATT 636
QY 207 PheProTyrLeuHisCysMetTyrPheHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCys 226
DB 637 ACCACAAATTCATGATGATGGTGGATATTTTAACTGGCCTGGTTCCTATCTTCAC--- 693
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLys 246
DB 694 ATCTCTTTCAGTTGTATACAAACAACTTTTACCTG---AGATATAGCCCAAAAGTGAAG 750
QY 247 Phe-----TrpPro 249
DB 751 TTTCTCTTGGAAATCTGGCCA 771
RESULT 12
ABZ111993
ID ABZ111993 standard; cDNA; 1063 BP.
XX
AC ABZ111993;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 875.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; neutropenic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RM, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2002-759812/82.
DR P-PSDB; ABP69776.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (SSTs), useful for treating cell-proliferative, or platelet
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 875; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP6849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical


```
US-10-017-410-4 (1-275) x AAA96499 (1-2744)
QY 14 SerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
DB 84 ACCTCCAGCTGGACTGGTGGAGAGAACTACTCCGTGACCTGGTACATGCCGAGTTC 143
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
DB 144 TGAATAACAGTGAAGTAAC---CTGATCATGATTATACCTCAATGTTGGTGGCAATT--- 197
QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuLeuValVal 73
DB 198 CAGAGTGTAGAGCGGTCTCGAAAAAGAGGACATGCTTCTTATTAGCACTCACAGT 257
QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
DB 258 GTAGGAATGGATCCTGGTGTCTTCCACATGACTCTGAATAATGAATGCCAGTATTGGAT 317
QY 94 GluLeuAlaValLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgArgTyr 113
DB 318 GAACCTCCCAATGATATAC---AGCTGTTGCATATTTTGTGTAAGTGTGTTGAATGTTTC 374
QY 114 LeuProIlyIlePheArgAsnAspArgGlyArgPheValValValValSerValLeuSer 133
DB 375 AAGATCAGAACTCAGTAACTACCACTGCTTTTACCTTAGTTCATTCAGTTTAAATA 434
QY 134 AlaValThrThrCysLeuAlaPheValIysProAlaIleAsnIleSerLeuMetThr 153
DB 435 GTAACACAGTTTACCTTAAGTAAAGAGCAATATTCATCAGGTCATGTAT----- 488
QY 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArg 173
DB 489 -----GGATGTTG 497
QY 174 ValPheLysLeuGlyLeuPheSer----- 181
DB 498 GTCTTTTACATAGTACTTCGATCTATTATATGTTTACATGGGTTTATCCATGGCTTAGA 557
QY 182 GlyLeuTyrTrpThr-----LeuAlaLeuPheCysTyrIleSerAsp 195
DB 558 GGACTGGTATATACATCATCTGGTATATTTTATTTGGATTTTATTTTGGAAATATAGAT 617
QY 196 ArgAlaPheCysGluLeuLeuSerSerPheAsn----- 206
DB 618 AACATATTTTGTGAGTCACTGAGAACCTTTCGAAAGAGGTACCACTCATCATAGTATT 677
QY 207 PheProTyrLeuHisCysMetTyrPheIleLeuIleCysLeuAlaAlaTyrLeuGlyCys 226
DB 678 ACCACAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLys 246
DB 735 ATCCTTTTCACTTTGTATACAGAACACTTTTACCTG---AGATATAGGCCAAAGTGAAG 791
QY 247 Phe-----TyrPro 249
DB 792 TTTCTCTTTGGAACTGGCCA 812
RESULT 14
AAC40178
ID AAC40178 standard; DNA; 1239 BP.
XX AAC40178;
AC AAC40178;
XX
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27344.
XX Arabidopsis thaliana.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126284P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132861P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140931P.
XX 30-JUN-1999; 99US-0141287P.
```



```
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 03-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142202P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-014814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1-12e-09
Score: 191.00
Percent Similarity: 42.05%
Best Local Similarity: 24.73%
Query Match: 12.71%
DB: 3
Length: 1239
Matches: 70
Conservative: 49
Mismatch: 107
Indels: 59
Gaps: 14

US-10-017-410-4 (1-275) x AAC40178 (1-1239)

QY 2 GlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCysGlu 21
Db 254 GGGATATCGAGCTTTTGGGTCCTGTG-----ACTTCTACTATAGAGTCTGTGAG 304
QY 22 AspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerIleValLeu 41
Db 305 ATGAACACTACGCTACTCATCTACATTTCTAGTCTCTACAACTATATCCCAATGTC--- 361
QY 42 PhePheIleLeuProPheIleCysMetCysLeuPheAspGluTyrAlaThrCysLeuAsn 61
Db 362 -----CCTGGAATCCTATTGGCTCTCATGGT---CTTGTCATGCAATTAGG 406
QY 62 SerAspIleTyrLeuIleTrpThrLeuLeu-----ValValValGlyIleGly 77
Db 407 CAACGGTTTGAGAGAGGTTTAGCATTTCTTCAATTTCAAAATATGATCTTGTCTATCGG 466
QY 78 SerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaVal 97
Db 467 AGCATGCTCTACCATGCCACCTTTTGAACACGTCGACACAGAGTGTGAGACCCCAATG 526
QY 98 LeuTrp---ValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLys 116
```


XX WPI; 2002-706409/76.
 DR P-PSDB; ABG32881.
 XX
 PT Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
 PT liver cancer, is differentially expressed in liver tumors relative to
 PT normal liver tissues.
 XX
 PS Claim 2; Page 8; 11pp; English.
 XX
 CC The invention relates to a polypeptide designated as CRG-L1, which is
 CC differentially expressed in liver tumors relative to normal expression
 CC in normal liver tissues, designated CRG-L1. Also included are the
 CC encoding polynucleotides (in the case of the human sequence, mapping to
 CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
 CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
 CC sequence in the diagnosis of hepatocellular cancer in tumour cells from a
 CC liver of a human or non-human animal. The CRG-L1 protein and
 CC polynucleotide are useful as diagnostic markers for a liver cancer in
 CC humans and non-human animals, and as a system for assessing putative
 CC therapeutic agents. The present sequence encodes human CRG-L1
 XX
 SQ Sequence 828 BP; 157 A; 216 C; 211 G; 244 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.98e-276 Length: 828
 Score: 275.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x ABS54149 (1-828)

Qy 1 MetGlyAlaProHisTrrTpAspGlnLeuGlnAlaGlySerSerGluValAspTrrCys 20
 Db 1 ATGGGCGCCCGACCTGGTGGGACAGCTGCAGGCTGGTAGCTCGGAGGTGGACTGGTGC 60

Qy 21 GluAspAntyrrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
 Db 61 GAGGACAACTACCATCGTCGCTATCCCGAGTTCTACACACGATCAGCAAGTGC 120

Qy 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
 Db 121 TTAATTTTCAATTTACGCCCATCTGCATGTGCTTTGTTGATGATGCAACATGCTTG 180

Qy 61 AsnSerAspIleTyrLeuIleTrrThrLeuLeuValValGlyIleGlySerValTyr 80
 Db 181 AACAGTGACATCTACTTAATCTGGACTCTTTTGGTTGAGTGGGAATTTGGATCCGCTAC 240

Qy 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrrVal 100
 Db 241 TTCATTTTACCTTTAGTTTCTTGGTTCAGATGCTTGTGATGAATTCGATGCTTTGGTT 300

Qy 101 LeuMetCysAlaLeuAlaMetTrrPheProArgGlyTrrLeuProLysIlePheArgAsn 120
 Db 301 CTGATGTGTGTTTGGCCATGTGGTCCCGAGAGGTATCTACCAAGATCTTTCCGAAT 360

Qy 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
 Db 361 GACAGGGGTAGTGTTCAGGTGGTGGTGTGCTGTGCTGTGGTTTACGACGTGCTGGCA 420

Qy 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
 Db 421 TTTGTCAAGCTCGCCATCAACACATCTCTGTGATGACCTCGGAGTTCTTTGCACTGCA 480

Qy 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
 Db 481 CTGCTCATCGCAGAGCTAAAGAGGTGTGACACATGCGTGTGTTTAAAGTGGCGCTCTTC 540

Qy 181 SerGlyLeuTrrThrIleAlaLeuPheCysTrrIleSerAspArgAlaPheCysGlu 200
 Db 541 TCGGGCCCTGTGGTGGACCTCGCCCTGTTTCTGCTGATGATGACCGAGCTTTCTGGAG 600

Qy 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrrPheIleLeuIleCysLeu 220
 Db 601 CTGCTGTCTCCTTCACTTCCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

Qy 221 AlaAlaTrrLeuGlyCysValCysPheAlaTyrPheAspAlaIleAsnGlyIleProGlu 240
 Db 661 GCTGCTTACCTGGGCTGTGTATGCTTTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

Qy 241 GlnGlyProValIleLysPheTrrProAsnGluLysTrrPheIleGlyValProTyr 260
 Db 721 CAAGGCGCTGTCTCAAGTTCTGCCCCTCAATGAGAAATGGGCTTTCATGTTGTTGCTTAT 780

Qy 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
 Db 781 GTGTCCCTCTGTGTGCTGCAACAGAAATCATCAGTCAAGATCAG 825

RESULT 2
 ACC50961
 ID ACC50961 standard; cDNA; 4202 BP.
 XX ACC50961;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human bladder cancer associated cDNA sequence SEQ ID NO:21.
 XX
 KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2003003906-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 03-JUL-2002; 2002WO-USO21338.
 XX
 PR 03-JUL-2001; 2001US-0302814P.
 PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Aziz N;
 XX
 DR WPI; 2003-201532/19.
 DR P-PSDB; ABR48156.
 XX
 PT Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX
 PS Claim 6; Page 235; 307pp; English.
 XX
 CC The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridises to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 SQ Sequence 4202 BP; 1045 A; 892 C; 927 G; 1338 T; 0 U; 0 Other;

Alignment Scores: 3.23e-189 Length: 4202
 Pred. No.: 192.00 Matches: 192
 Score: 100.00% Conserv: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 69.82% Gaps: 0
 DB: 7

US-10-017-410-4 (1-275) x ACC50961 (1-4202)

QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103
 DB 326 ACCCTTAGTTCTTGGGTGAGATGCTGATGAACCTTGCAGTCTTGGGTCTGATGCT 385

QY 104 AlaLeuAlaMetTrpPheProArgTrpLeuProLysIlePheArgAsnAspArgGly 123
 DB 386 GCTTGGCAUGTGTCTCCCAAGAGGTATCTACCAAGATCTTTCGGGAATGACCGGGT 445

QY 124 ArgPheLysValValSerValLeuSerAlaValThrCysLeuAlaPheValLys 143
 DB 446 AGGTTCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505

QY 144 ProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle 163
 DB 506 CTGCGCATCAACATCTCTGATGACCCCTGGGAGTTCTTGCACCTGCATGCTCATC 565

QY 164 AlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeu 183
 DB 566 GCAGAGCTAAGAGGTGTGACAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 625

QY 184 TrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSer 203
 DB 626 TGGTGCACCTGGCCCTGTCTGCTGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCA 685

QY 204 SerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaIleTyr 223
 DB 686 TCCTTCAACTTCCCTACTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 745

QY 224 LeuGlyCysValCysPheAlaIleTyrPheAspAlaIleSerGluIleProGluGlnGlyPro 243
 DB 746 CTGGCTGTGTATGTTTGGCTTACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 805

QY 244 ValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
 DB 806 GTCATCAAGTTCTGGCCCAATGAGAAATGGGCTTCATTTGTTGTTCCCTATGTGTCCTC 865

QY 264 LeuCysAlaAsnLysLysSerSerValLysIleThr 275
 DB 866 CTGTGTGCCAACAGAAATCATCAGTCAAGATCAGC 901

RESULT 3
 AAH48641
 ID AAH48641 standard; DNA; 822 BP.
 XX AC AAH48641;
 XX 21-SEP-2001 (first entry)
 XX Human ceramidase K3 DNA.
 KW Ceramidase; human; K3; antiproliferative; anticancer; anti-eczema;
 KW antipsoriasis; dermatological; ceramide; sphingosine; treatment; ds;
 KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
 KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.
 XX Homo sapiens.
 OS
 XX Key
 FH Location/Qualifiers
 FT 1. 822
 CDS
 FT /*tag= a
 FT /transl_except= (pos:382..387,aa:Cys)
 FT /note= "No stop codon given"

XX WO200155408-A1.
 XX 02-AUG-2001.
 PD 27-JAN-2001; 2001WO-EP000900.
 PF 27-JAN-2000; 2000DE-01003293.
 XX 09-MAR-2000; 2000DE-01011392.
 PR
 PR (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.
 PA Hofmann K, Conradt M;
 XX WPI: 2001-483256/52.
 DR P-PSDB; AAB86365.
 XX New ceramidase containing specific structural motif, useful for
 PT diagnosis, prevention and treatment of ceramidase defects, e.g.
 PT ichthyosis, also in cosmetics.
 XX Claim 3; Page 17; 31pp; German.
 CC This invention describes novel human ceramidase (I) containing a specific
 CC structural motif which has antiproliferative, anticancer, anti-eczema,
 CC antipsoriasis and dermatological activity. (I) cleave ceramide, resulting
 CC in formation of sphingosine (I), or the nucleic acid (II) encoding it,
 CC are used for diagnosis, prevention or treatment of diseases associated
 CC with ceramide defects, particularly altered cell proliferation (cancer)
 CC or altered ceramide layers on the skin (neurodermatitis, eczema,
 CC psoriasis), also for targeted modification of the permeability barrier by
 CC ceramidase or its activators, e.g. for transcutaneous delivery of
 CC substances. Specifically they are used for diagnosis of ichthyosis,
 CC particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
 CC This sequence encodes human ceramidase K3
 XX SQ Sequence 822 BP; 153 A; 225 C; 209 G; 235 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.47e-78 Length: 822
 Score: 85.00 Matches: 85
 Percent Similarity: 100.00% Conserv: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 30.91% Indels: 0
 DB: 4 Gaps: 0
 US-10-017-410-4 (1-275) x AAH48641 (1-822)
 QY 165 GluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrp 184
 DB 490 GAGCTAAGAGGTGTGACAAACATGGTGTGTTAAGCTGGGCCCTCTTCTGGGCCCTCTGG 549
 QY 185 TrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSer 204
 DB 550 TGGACCTTGGCCCTGTTCTGCTGGATCAGTACCGAGCTTTCTGCGAGCTGTGTCTATCC 609
 QY 205 PheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaIleTyrLeu 224
 DB 610 TTCACCTTCCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
 QY 225 GlyCysValCysPheAlaIleTyrPheAspAlaIleSerGluIleProGluGlnGlyProVal 244
 DB 670 GGCTGTGTATGCTTGTGCTTACTTGTATGCTGCTCAGAGATTCCTGACCAAGGCCCTGTC 729
 QY 245 IleLysPheTrpPro 249
 DB 730 ATCAAAATTCGGCC 744
 RESULT 4
 ABS54148
 ID ABS54148 standard; cDNA; 4175 BP.
 XX AC ABS54148;

```
XX 25-NOV-2002 (first entry)
DT Mouse cDNA encoding a liver tumour marker protein, CRG-L1.
XX
DE Mouse cDNA encoding a liver tumour marker protein, CRG-L1.
XX
KW Mouse; ss; gene; liver cancer; liver tumour; CRG-L1;
KW hepatocellular cancer.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
FH 35..862
FT /*tag= a
FT /product= "Liver tumour marker protein, CRG-L1"
FT
XX
XX US2002115094-A1.
XX
XX 22-AUG-2002.
XX
XX 14-DEC-2001; 2001US-00017410.
XX
XX 14-DEC-2000; 2000US-0255674P.
XX
XX (PARN/) FARNHAM P J.
XX (GRAV/) GRAVEEL C R.
XX
XX Farnham PJ, Graveel CR;
XX
XX WPI; 2002-706409/76.
XX P-PSDB; ABG32880.
XX
XX Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
PT liver cancer, is differentially expressed in liver tumors relative to
PT normal liver tissues.
PT
XX Claim 2; Page 4-7; 11pp; English.
XX
XX The invention relates to a polypeptide designated as CRG-L1, which is
CC differentially expressed in liver tumours relative to normal expression
CC in normal liver tissues, designated CRG-L1. Also included are the
CC encoding polynucleotides (in the case of the human sequence, mapping to
CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
CC sequence in the diagnosis of hepatocellular cancer in tumour cells from a
CC liver of a human or non-human animal. The CRG-L1 protein and
CC polynucleotide are useful as diagnostic markers for a liver cancer in
CC humans and non-human animals, and as a system for assessing putative
CC therapeutic agents. The present sequence encodes mouse CRG-L1
XX
SQ Sequence 4175 BP; 962 A; 971 C; 1012 G; 1230 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 3,82e-33 Length: 4175
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.27% Indels: 0
DB: 6 Gaps:
```

US-10-017-410-4 (1-275) x ABS54148 (1-4175)

```
Qy 12 AlaGlySerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAla 31
Db 68 GCTGGCAGTTCGGAGGTGGATTGGTGGAGGACAACTACATCGTGGCTGCATTGCC 127
Qy 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProPheIleCysMetCys 51
Db 128 GAGTCTACACACATCAGCAACGCTGTGTTTTCATTTTACCTCCCATCTGCATGTC 187
Qy 52 LeuPhe 53
Db 188 TTGTC 193
```

```
RESULT 5
AAC40178
ID AAC40178 standard; DNA; 1239 BP.
XX
XX AAC40178;
AC
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27344.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 09-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 08-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 08-MAY-1999; 99US-0132485P.
XX 08-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
```

PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145911P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 23-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160988P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 0.216 Length: 1239
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.00% Indels: 0
 DB: 3 Gaps: 0

US-10-017-410-4 (1-275) x AAC40178 (1-1239)

Oy 30 IleAlaGluPheTyrAsnThrIleSerAsnVal 40

Db 329 ATTGCTGAGTTCACACACTATATCCATGTC 361

RESULT 6

AAT83805


```
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234224P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234977P.
25-SEP-2000; 2000US-0234988P.
25-SEP-2000; 2000US-0234989P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0241826P.
08-NOV-2000; 2000US-0246417P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.

106-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251899P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 37365; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 1584 BP; 335 A; 435 C; 464 G; 349 T; 0 U; 1 Other;

Alignment Scores:
Pred. NO.: 32.9 Length: 1584
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x AAK82553 (1-1584)
Qy 200 GluLeuLeuSerSerPheAsnPhePro 208
Db 714 GAGCTTCTCAGCTCCTTTAAATTTCC 688

RESULT 8
AAK82554/C
ID AAK82554 standard; DNA; 3279 BP.
XX
XX AAK82554;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37366.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
```


PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 37366; 3071pp + Sequence Listing; English.
PS
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 3279 BP; 828 A; 802 C; 914 G; 735 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 65.8 Length: 3279
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x AAK82554 (1-3279)

Qy 200 GluLeuLeuSerSerPheAsnPhePro 208
Db 713 GAGCTTCAGCTCCTTAATTTTCCC 687

RESULT 9

AAK82555/c

ID AAK82555 standard; DNA; 3280 BP.

XX AAK82555;

XX 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37367.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

Alignment Scores:
 Pred. No.: 854 Length: 48012
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.27% Indels: 0
 DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x ABS79084 (1-48012)

Qy 142 VallysProAlalleAsnAsnIleSer 150
 DB 35109 GTAAACACGCTATAAATACATTCA 35135
 |||||

RESULT 11

ABA89141
 ID ABA89141 standard; DNA; 48254 BP.

XX AC ABA89141;

DT 11-FEB-2002 (first entry)

DE Escherichia coli polynucleotide SEQ ID NO 829.

XX Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicemia;
 KW pyelonephritis; antibiotic resistance; ds.

XX Escherichia coli.

OS WO200166572-A2.

PN 13-SEP-2001.

PD 12-MAR-2001; 2001WO-EP003445.

XX 10-MAR-2000; 2000FR-00003145.

PR 02-FEB-2001; 2001FR-00001449.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;

XX WPI; 2001-550253/61.

PT A library of DNA fragments of Escherichia coli strains for the phylogenic
 PT determination of a given strain comprises polynucleotides of nature B2/D+
 A-.

XX Example 6; Fig 6; 646pp; English.

XX The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533) and
 CC encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli
 CC infections. The polypeptides are useful for determining the phylogenic
 CC group of a given E. coli strain. These polypeptides can detect and treat
 CC an undesired development of E. coli, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more frequent
 CC use of broad spectrum antibiotics

SQ Sequence 48254 BP; 12342 A; 11284 C; 12417 G; 12211 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 859 Length: 48254
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.27% Indels: 0
 DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x ABA89141 (1-48254)

Qy 142 VallysProAlalleAsnAsnIleSer 150
 DB 35424 GTAAACACGCTATAAATACATTCA 35450
 |||||

RESULT 12

ABA89142
 ID ABA89142 standard; DNA; 48345 BP.

XX AC ABA89142;

DT 11-FEB-2002 (first entry)

DE Escherichia coli polynucleotide SEQ ID NO 830.

XX Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicemia;
 KW pyelonephritis; antibiotic resistance; ds.

XX Escherichia coli.

OS WO200166572-A2.

PN 13-SEP-2001.

PD 12-MAR-2001; 2001WO-EP003445.

XX 10-MAR-2000; 2000FR-00003145.

PR 02-FEB-2001; 2001FR-00001449.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;

XX WPI; 2001-550253/61.

PT A library of DNA fragments of Escherichia coli strains for the phylogenic
 PT determination of a given strain comprises polynucleotides of nature B2/D+
 A-.

XX Example 6; Fig 6; 646pp; English.

XX The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533) and
 CC encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli
 CC infections. The polypeptides are useful for determining the phylogenic
 CC group of a given E. coli strain. These polypeptides can detect and treat
 CC an undesired development of E. coli, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more frequent
 CC use of broad spectrum antibiotics

SQ Sequence 48345 BP; 12347 A; 11290 C; 12423 G; 12219 T; 0 U; 66 Other;

Alignment Scores:

Pred. No.: 860 Length: 48345
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.27% Indels: 0
 DB: 4 Gaps: 0

US-10-017-410-4 (1-275) x ABA89142 (1-48345)

QY 142 ValLysProAlaIleAsnAsnIleSer 150
 Db 35424 GTAAACACGCTATAAATAACATTCA 35450

RESULT 13

AAF22305_08/c
 Continuation (9 of 11) of AAF22305 from base 800001 (Arabidopsis thaliana chromosome 4)
 WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305
 WP Fragment Name Begin End
 WP AAF22305_00 1 110000
 WP AAF22305_01 100001 210000
 WP AAF22305_02 200001 310000
 WP AAF22305_03 300001 410000
 WP AAF22305_04 400001 510000
 WP AAF22305_05 500001 610000
 WP AAF22305_06 600001 710000
 WP AAF22305_07 700001 810000
 WP AAF22305_08 800001 910000
 WP AAF22305_09 900001 1010000
 WP AAF22305_10 1000001 1082138

Alignment Scores:
 Pred. No.: 1.89e+03 Length: 110000
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.27% Indels: 0
 DB: 3 Gaps: 0

US-10-017-410-4 (1-275) x AAF22305_08 (1-110000)

QY 80 TyrPheHisPheThrLeuSerPheLeu 88
 Db 75906 TATTTTCATTTTACTTTGAGTTCTTA 75880

RESULT 14

ABN71527_17
 Continuation (18 of 22) of ABN71527 from base 1700001 (Streptococcus polynucleotide SEQ)
 WP Sequence split into 22 fragments LOCUS ABN71527 Accession Abn71527
 WP Fragment Name Begin End
 WP ABN71527_00 1 110000
 WP ABN71527_01 100001 210000
 WP ABN71527_02 200001 310000
 WP ABN71527_03 300001 410000
 WP ABN71527_04 400001 510000
 WP ABN71527_05 500001 610000
 WP ABN71527_06 600001 710000
 WP ABN71527_07 700001 810000
 WP ABN71527_08 800001 910000
 WP ABN71527_09 900001 1010000
 WP ABN71527_10 1000001 1110000
 WP ABN71527_11 1100001 1210000
 WP ABN71527_12 1200001 1310000
 WP ABN71527_13 1300001 1410000
 WP ABN71527_14 1400001 1510000
 WP ABN71527_15 1500001 1610000
 WP ABN71527_16 1600001 1710000
 WP ABN71527_17 1700001 1810000
 WP ABN71527_18 1800001 1910000
 WP ABN71527_19 1900001 2010000
 WP ABN71527_20 2000001 2110000
 WP ABN71527_21 2100001 2155561

Alignment Scores:
 Pred. No.: 1.89e+03 Length: 110000
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.27% Indels: 0
 DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x ABN71527_17 (1-110000)

QY 202 LeuSerSerPheAsnPheProTyrLeu 210
 Db 21744 CTTTCATCATTTAACTTCCTTACTTG 21770

RESULT 15

AAF67001
 ID AAF67001 standard; cDNA; 369 BP.
 XX
 AC AAF67001;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE Novel human polynucleotide, SEQ ID NO: 2757.
 XX
 KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200102568-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US018374.
 XX
 PR 02-JUL-1999; 99US-0142310P.
 PR 02-JUL-1999; 99US-0142311P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;
 PI Kassam A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
 PI Drmanac R, Crkenjakov R, Drmanac S, Dickson M, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
 XX
 DR WPI; 2001-091805/10.
 XX
 PT Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences.
 XX
 PS Claim 9; Page 953-954; 1046pp; English.

The present sequence is one of 3351 sequences in a library of human polynucleotides. The library is used to detect differentially expressed genes correlated with a cancerous state of a mammalian cell and can detect colon, prostate, breast and lung cancer. The library can be used to produce probes for detection of mRNA and to produce additional copies of the polynucleotides. The probes can be used for chromosome mapping of the polynucleotide and for detection of transcription levels. Ribozymes or antisense oligonucleotides can be generated. The polynucleotides and their gene products are used as genetic or biochemical markers (e.g. in blood or tissues) that will detect the earliest changes along the carcinogenesis pathway and/or monitor the efficacy of therapies and preventive interventions. The polynucleotides, polypeptides and antibodies against them can be used in pharmaceutical compositions to treat the cancers and proliferative disorders such as neoplasia, dysplasia and hyperplasia

SQ Sequence 369 BP; 83 A; 89 C; 87 G; 110 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 89.7 Length: 369
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.91% Indels: 0
 DB: 5 Gaps: 0

US-10-017-410-4 (1-275) x AAF67001 (1-369)

QY 176 LysLeuGlyLeuPheSerGlyLeu 183

Db 294 AACTAGGATTGTTCTCAGGCTT 317
|||||

Search completed: September 18, 2004, 07:13:35
Job time : 454 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 05:02:44 ; Search time 3210 Seconds

(without alignments)
3713.188 Million cell updates/sec

Title: US-10-017-410-4

Perfect score: 1503

Sequence: 1 MGAPHWQDLQAGSSEVDMC.....IGVPYVLLCANKKSSVKIT 275

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp

-Q=/cgn2_1/USPRO.spool_p/US10017410/runat_15092004_164709_4691/app_query.fasta_1.455

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10017410@cgn_1_1_5600@runat_15092004_164709_4691 -NCPU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pin.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1466	97.5	2852	9	AK123581	AK123581 Homo sapi
2	1404	93.4	4174	10	AF282864	AF282864 Mus muscu
3	1390.5	92.5	822	6	AX206799	AX206799 Sequence
4	1378.5	91.7	149462	9	AC017081	AC017081 Homo sapi
5	1169	77.8	4078	10	BC059819	BC059819 Mus muscu
6	793.5	52.8	1429	3	AK116177	AK116177 Ciona int
7	714.5	47.5	1163	3	AF323976	AF323976 Drosophil
8	709.5	47.2	2318	3	AY071232	AY071232 Drosophil
9	538.5	35.8	792	6	AX206797	AX206797 Sequence
10	538.5	35.8	795	9	AF347024	AF347024 Homo sapi
11	517.5	34.4	822	10	AF347023	AF347023 Mus muscu
12	335.5	22.3	163542	9	AL158206	AL158206 Human DNA
13	335.5	22.3	219679	2	AC109451	AC109451 Homo sapi
14	319	21.2	184163	2	AC120245	AC120245 Rattus no
15	319	21.2	249734	2	AC097362	AC097362 Rattus no
16	316	21.0	90370	10	BX005084	BX005084 Mouse DNA
17	316	21.0	207360	2	BX324222	BX324222 Mus muscu
18	308.5	20.5	112713	9	AL391834	AL391834 Human DNA
19	308.5	20.5	219679	2	AC109451	AC109451 Homo sapi
20	307	20.4	13715	2	AC017332	AC017332 Drosophil
21	307	20.4	78857	3	AC004364	AC004364 Drosophil
22	307	20.4	157875	3	AE003665	AE003665 Drosophil
23	307	20.4	164361	3	AC009252	AC009252 Drosophil
24	307	20.4	165585	3	AC093044	AC093044 Drosophil
25	282.5	18.8	38138	2	AC017406	AC017406 Drosophil
26	216.5	14.4	801	6	AX206798	AX206798 Sequence
27	216.5	14.4	3404	9	AF214454	AF214454 Homo sapi
28	216.5	14.4	3473	9	AF327353	AF327353 Homo sapi
29	213	14.2	358	6	AR419744	AR419744 Sequence
30	213	14.2	358	6	BD115297	BD115297 EST and e
31	202.5	13.5	1460	10	BC023924	BC023924 Mus muscu
32	201	13.4	954	8	AF214455	AF214455 Saccharom
33	201	13.4	38669	8	SCU43281	U43281 Saccharomyc
34	198.5	13.2	1313	8	AK103953	AK103953 Oryza sat
35	198.5	13.2	1335	8	AK073303	AK073303 Oryza sat
36	196	13.0	799	8	BT008549	BT008549 Arabidops
37	196	13.0	1193	8	AY090947	AY090947 Arabidops
38	196	13.0	1204	8	BT008652	BT008652 Arabidops
39	196	13.0	1235	8	AB063253	AB063253 Arabidops
40	186	12.4	951	8	AF191745	AF191745 Saccharom
41	186	12.4	2498	8	SCYBR183W	Z36052 S.cerevisia
42	186	12.4	12800	8	SCC1IMJB	U02073 Saccharomyc
43	186	12.4	35425	6	AX647877	AX647877 Sequence
44	186	12.4	118820	9	AC011491	AC011491 Homo sapi
45	184.5	12.3	2146	9	BC063034	BC063034 Homo sapi

ALIGNMENTS

RESULT 1

```

AK123581
LOCUS      Homo sapiens cDNA FLJ41587 fis, clone CTONG2020638.
DEFINITION
ACCESSION  AK123581
VERSION    AK123581.1 GI:34529163
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Osuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahara,K., Maeuho,Y., Nagai,K. and Isogai,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 2852)
            Isogai,T. and Yamamoto,J.
            Direct Submission
            Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 252-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
            Location/Qualifiers
            1..2852
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CTONG2020638"
            /tissue type="tongue, tumor tissue"
            /clone lib="CTONG2"
            /note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.:      1..89e-132      Length:      2852
Score:          1466.00      Matches:      269
Percent Similarity: 98.18%      Conservative: 1
Best Local Similarity: 97.82%      Mismatches: 5
Query Match:     97.54%      Indels:      0
DB:              9      Gaps:          0

US-10-017-410-4 (1-275) x AK123581 (1-2852)

QY      1  MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db      159  ATGGCGCCCGCCACCTGGGGGACCGAGCTGGAGTGGTAGTCGCGAGGTGGAGTGGTGC 218
QY      21  GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db      219  GAGGACAACTACACCATCGTGCCTGCTATCGCCGAGTCTCAACACGATCAGCATGTC 278
QY      41  LeuPhePheIleLeuProProlleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db      279  TTATTTTTCATTATCCCGCCCATCTGCATGCTGTTGTCGTCAGTATCAACATGCTTC 338
QY      61  AsnSerAspIleTyrIleIleTrpThrIleLeuValValValGlyIleCysValTyr 80
Db      339  AACAGTGGCATCTACTTAATCGGACTCTTTGGTTGTAGTGGGAATGGAATCCGCTAC 398
QY      81  PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100

```

```

Db      399  TTCCATGCAACCTTAGTTTCTTGGGTGAGATGCTTTGATGAACCTTCGACTCTTTGGGTT 458
QY      101  LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db      459  CTGATGTGTCTTTGGCCATGTGGTTCCCGAAGGATATCTACCAAGATCTTTTGGAAAT 518
QY      121  AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db      519  GACCGGGGTAGGTTCAAGGTGGTGGTCACTGTCTGTCTGCGGTACGACGAGTCCCGGCA 578
QY      141  PheValLysProAlaIleAsnAsnIleSerLeuMetThrIleGlyValProCysThrAla 160
Db      579  TTTTGCAAGCCCTGCCATCAACAACATCTCTCTGATGACCTCGGAGTCTCTTGCACGTCA 638
QY      161  LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db      639  CTGCTCATCGCAGAGTAAGAGGTGTGACACATGCGTGTGTTAAGGTGGGCTCTTTC 698
QY      181  SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTyrPheSerAspArgAlaPheCysGlu 200
Db      699  TCGGGCTCTGTGGACCTGGCCCTGTTCTGTGATCAGTCAGTCAGCGAGCTTCTCGGAG 758
QY      201  LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db      759  CTGCTGTCTATCTTCAACTTCCCTACCTGCATGTCATGTCACATCTCATCTGCTCTT 818
QY      221  AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db      819  GTTGCTACTTGGGTGTGTATGCTTTGCTACTTTGATGCTGCTCCTCAGAGATTCCTGAG 878
QY      241  GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyr 260
Db      879  CAAGGCCCTGTCTCAAGTTCCTGGCCCATGAGAAATGGCTTCATCTGTTGCTCCCTAT 938
QY      261  ValSerLeuLeuCysAlaAsnLysSerSerValLysIleThr 275
Db      939  GTGTCCCTCTGTGTGCCAACAGAAATCATCAGTCAAGATCAGC 983

RESULT 2
AF282864      4174 bp      mRNA      linear      ROD 12-DEC-2001
Mus musculus cancer related gene-liver 1 mRNA, complete cds.
AF282864
AF282864.1      GI:17529683
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 4174)
AUTHORS       Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
TITLE         Expression profiling and identification of novel genes in
              hepatocellular carcinomas
JOURNAL       Oncogene 20 (21), 2704-2712 (2001)
MEDLINE       21313787
PUBMED        11420682
REFERENCE     2 (bases 1 to 4174)
AUTHORS       Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
TITLE         Identification of genes deregulated in murine hepatocellular
              carcinomas using oligonucleotide microarrays and representational
              difference analysis
JOURNAL       Unpublished
REFERENCE     3 (bases 1 to 4174)
AUTHORS       Farnham,P.J. and Graveel,C.R.
TITLE         Direct Submission
JOURNAL       Submitted (27-JUN-2000) Oncology, University of Wisconsin, 1400
              University Avenue, Madison, WI 53706, USA
              Location/Qualifiers
              1..4174
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /tissue_type="hepatocellular carcinoma"

```

CDS

35...862
/note="CRG-L1: putative"
/codon_start=1
/product="cancer related gene-liver 1"
/protein_id="AAL40408.1"
/db_xref="GI:17529684"
/translation="MGAPHWMDHLRAGSSVDCEQNTYIVPAIAPYNTISNLFPI
LPICWCLPQVATCFNSGYLITWLLVVGIGSVYFHTLSFLGQMLDLAILWLIM
CALAMFPBRYLKIIFNDGRPKAVCVLSAITCLAFIKPAININISLMIGPCTA
LLVAELKRCNVRVFKLGLPSGLWNTLALGKISDQAPCELLSSFPHPYLHCWHLI
CLASYLGVCVCFVDAASEIPEQGPVIRFPWSEKWFIFGVYVSLLCIAHKKSPVKIT"

ORIGIN

Alignment Scores:
Pred. No.: 3 13e-126 Length: 4174
Score: 1404.00 Matches: 251
Percent Similarity: 96.73% Conservative: 15
Best Local Similarity: 91.27% Mismatches: 9
Query Match: 93.41% Indels: 0
DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x AF282864 (1-4174)

Qy 1 MetGlyAlaProHisTrpTyrAspGlnLeuGlnAlaGlySerSerGluValAspTyrCys 20
Db 35 ATGGGCGCCCGCAGTGGTGGGACCACTGGGGCTGGCAGTTCGGAGGTGGATTGGTGC 94
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 95 GAGGACAACTACACTATCGTGGCTGCCATTGCCAGTTCACACAGCATCAGCAAGCTC 154
Qy 41 LeuPhePheIleLeuProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db 155 TTGTTTTCATTACCTCCCATCTGATGCTGTTCGCCAGTACGCAAGTGTTC 214
Qy 61 AsnSerAspIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
Db 215 AACAGCGCATCTACTTAATATGACGCTCCTAGTGTAGTGGGATGGATCTGTCTAC 274
Qy 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTyrVal 100
Db 275 TTCCATGCAACCGTGGATTTCTGGGTGAGTGTTCGATGAATTCGCAATCTCTGGGTT 334
Qy 101 LeuMetCysAlaLeuAlaMetTyrPheProArgTyrLeuProLysIlePheArgAsn 120
Db 335 CTGATGTGTCTTGGCCATGTGGTTTCCAGAGAGTATTACCAAGATCTTTCGGAAT 394
Qy 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 395 GACAGGGGAGGTTCAAGGCGAGTGTGTCTGTCTGTCTGCAATTAACAACGTGTGGCG 454
Qy 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 455 TTTATCAAGCCCGCCATCAACAATTTCCCTGATGATTTCTGGGACTTCCATGCACTGCG 514
Qy 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db 515 CTGTTTGTGGAGCTGAAGAGTGTGACATGTGGGTGTGTATAGCTGGGCTCTTC 574
Qy 181 SerGlyLeuTyrTrpThrLeuAlaLeuPheCysTrpIleSerAspAlaPheCysGlu 200
Db 575 TCTGGCCTCTGGTGGACTCTGCTCTCTGCTGGATCAGCGACCAAGCCCTTCTGTGAG 634
Qy 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTyrHisIleLeuIleCysLeu 220
Db 635 CTGCTCTCTCTTTCATCTCCCTTACCTGCACTGTGTGTGGCATATTCTCATCTGCGCT 694
Qy 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 695 GCTTCGTAACCTGGGCTGTGTGCTTCCCTACTTTGATGCTGCTCAGATACCTGAG 754
Qy 241 GlnGlyProValIleLysPheThrProAsnGlnLysTyrAlaPheIleGlyValProTyr 260

Db 755 CAAGGTCCAGTCATCAGATTCTGGCCAGCGAAGATGGGCTTTATTGGTGTCCCTTAT 814
Qy 261 ValSerLeuLeuCysAlaAsnLysSerSerValLysIleThr 275
Db 815 GTGTCCCTTCTGTGTGCCCAAGAGATGCGCCAGTCAAGATCAG 859

RESULT 3

LOCUS AX206799 822 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 6 from Patent WO0155408.
ACCESSION AX206799
VERSION AX206799.1 GI:15394643
KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .822

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 9 04e-126 Length: 822

Score: 1390.50 Matches: 256

Percent Similarity: 94.55% Conservative: 4

Best Local Similarity: 93.09% Mismatches: 14

Query Match: 92.51% Indels: 1

DB: 6 Gaps: 1

US-10-017-410-4 (1-275) x AX206799 (1-822)

Qy 1 MetGlyAlaProHisTrpTyrAspGlnLeuGlnAlaGlySerSerGluValAspTyrCys 20
Db 1 ATGGGCGCCCGCAGTGGTGGGACCACTGGGGCTGGCAGTTCGGAGGTGGATGGCGC 60
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 61 GAGGACAACTACACTATCGTGGCTGGTGTCTGCGCAGTTCCTATAACATCAGCAATGTC 120
Qy 41 LeuPhePheIleLeuProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db 121 TTATTTTTCATTTTACCGCCCATCTGCATGTGTCTGTCTGTCAGTATGCAACATGCTTC 180
Qy 61 AsnSerAspIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
Db 181 AACAGCGCATCTACTTAATCTGG---CTCTGTGGTTAGCGGAATGGATCGCTCTAC 237
Qy 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTyrVal 100
Db 238 TTCCATGCAACCTTAGTTTCTGGGTGAGTGTGATGATGAGTTCAGTCTCTTGGGTT 297
Qy 101 LeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db 298 CTGATGTGTGCTTCCGTGATGTGGTGTCCCGAAGAGTATCTACCAAGATCTTTCGGAAT 357
Qy 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 358 GACAGGGTAGGTTCAGAGGTGGTGTGTCTGTCTGTCTGTCAGTATGACAGTGTGGCA 417
Qy 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 418 TTTGTCAAGCTGCGCATCAACACATCTCTGATGACCTGGGAGTTCCTTGGCTGCA 477
Qy 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180

Db 478 CTGCTCATCACAGAGCTAAGAGGTGTGACAAACATCGGTGTGTTTAAAGTGGCCCTCTTC 537
Qy 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db 538 TGGGGCTCTGGTGGACCTGGCCCTGTCTGTGTGATCAGTACCGAGCTTCTCGAG 597
Qy 201 LeuLeuSerSerPheAsnPheProTrpLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db 598 CTGCTGTCATCTTCAACTCCCTACCTGACATGCTGCGACATGCGACATCTCATCTGCCIT 657
Qy 221 AlaAlaTyrlendlyCysValCysPheAlaTyrlPheAspAlaAlaSerGluIleProGlu 240
Db 658 GCTGCTACCTGGCTGTGTATCTTGTGCTACTTTCATGCTGCTCCTCAGAGATTCCTGAG 717
Qy 241 GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTrp 260
Db 718 CAAGGCCCTGTCAAAATTCGGCCAGGAGAAATGGCCCTTCATGTGTGTCCTCAT 777
Qy 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
Db 778 GTGTCCCTCTGTGTGCCAACAGAAATCATCATGTCAGACACG 822

RESULT 4

AC017081/c
LOCUS AC017081 149462 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-470J24 from 2, complete sequence.
ACCESSION AC017081
VERSION AC017081.8 GI:18093316

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 149462)

Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

99063792

9847074

2 (bases 1 to 149462)

Nguyen, C., Doebber, A. and Kozlowski, A.

The sequence of Homo sapiens BAC clone RP11-470J24

Unpublished (2001)

3 (bases 1 to 149462)

Waterston, R.H.

Direct Submission

Submitted (09-DEC-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 149462)

Waterston, R.

Direct Submission

Submitted (09-JAN-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Jan 9, 2002 this sequence version replaced gi:14165368.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>Contact: sapiens@wustl.wustl.edu

----- Summary Statistics

Center project name: H_NH0470J24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Piatier de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-310K15, 2000 bp overlap; the clone sequenced to the right is RP11-156A1, 2000 bp overlap. Actual start of this clone is at base position 190775 of RP11-310K15; actual end is at base position 28935 of RP11-156A1.

The sequence between 81601 and 81703 is covered only by a PCR product of clone DNA. Unresolved tandem repeat regions exist between 81548 and 83183, 126196 and 127426.

FEATURES

source

Location/Qualifiers
1. .149462
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-470J24"
/clone_lib="RPC1-11"
141. .282
/rpt_family="MIR"
581. .621
/rpt_family="TGGG)n"
965. .1037
/rpt_family="GC_rich"
1127. .1561
/note="match to EST AA056210 (NTD:G1548612) zf62g06.s1"
1771. .2127
/note="match to EST AW614820 (NTD:G7320006) hg80c08.x1"
2055. .2195
/rpt_family="MER1_type"
2216. .2419
/rpt_family="MER1_type"
2604. .2844
/rpt_family="MIR"
2991. .3324
/note="match to EST R35058 (NTD:G791959) yh86d09.r1"
3207. .3758
/note="match to EST AW452004 (NTD:G6992780)"
3486. .3527
/rpt_family="AT_rich"
3501. .3745
/note="match to EST AA054208 (NTD:G1545131) zf51g03.r1"
3981. .4017
/rpt_family="T-rich"
4231. .4273
/rpt_family="AT_rich"
5518. .5847
/note="match to EST R80994 (NTD:G857275) yi94c08.r1"
5606. .5633
/rpt_family="AT_rich"
5722. .5758
/rpt_family="AT_rich"
5802. .5831

BC059819 4078 bp mRNA linear ROD 21-OCT-2003
 LOCUS Mus musculus cancer related gene-liver 1, mRNA (cdna clone
 DEFINITION MGC:69583 IMAGE:6839525), complete cds.
 ACCESSION BC059819
 VERSION BC059819.1 GI:37590519
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4078)
 Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zerbino,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,J.,
 Stapleton,M., Soares,M.B., Bonaldo,A.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalski,U., Smalins,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A., M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 4078)
 Strausberg,R.
 Direct Submission
 Submitted (07-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-x@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lini.gov>
 Series: IRAC Plate: 132 Row: 0 Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21314857.

Location/Qualifiers
 1..4078
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:69583 IMAGE:6839525"
 /tissue_type="Brain"

FEATURES
 source

gene
 /clone_lib="NIH_BMAP_GH0"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1..4078
 /genes="CRG-L1"
 /db_xref="LocusID:230379"
 76..765
 /codon_start=1
 /product="CRG-L1 protein"
 /protein_id="AAH59819.1"
 /db_xref="GI:37590520"
 /db_xref="LocusID:230379"
 /translation="MGAPHWMDHLRASGSSEVDMCEDNYITVPAIABFYNTISNVLFFI
 LPIICMLFRQATYSGISGTYITWLVVVGISGVYFATLSFLGQMLDELAILWLM
 CALAMFPRYLPKIPRNDRCNVRVFKLGFSLWTLALFCWISDOAPCELLSFFH
 EYVLCVWHILICLASLYLGCVCFAFEDAASEIPEQGPVIRFWSEKWARFIVGVYSLL
 CHAKSKSPVKIT"

CDS
 /codon_start=1
 /product="CRG-L1 protein"
 /protein_id="AAH59819.1"
 /db_xref="GI:37590520"
 /db_xref="LocusID:230379"
 /translation="MGAPHWMDHLRASGSSEVDMCEDNYITVPAIABFYNTISNVLFFI
 LPIICMLFRQATYSGISGTYITWLVVVGISGVYFATLSFLGQMLDELAILWLM
 CALAMFPRYLPKIPRNDRCNVRVFKLGFSLWTLALFCWISDOAPCELLSFFH
 EYVLCVWHILICLASLYLGCVCFAFEDAASEIPEQGPVIRFWSEKWARFIVGVYSLL
 CHAKSKSPVKIT"

ORIGIN
 Alignment Scores: 2e-103 Length: 4078
 Pred. No.: 1169.00 Matches: 212
 Score: 81.09% Conservative: 11
 Percent Similarity: 77.09% Mismatches: 6
 Best Local Similarity: 77.78% Indels: 46
 Query Match: 10 Gaps: 1
 DB: 10
 US-10-017-410-4 (1-275) x BC059819 (1-4078)

QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
 Db 76 ATGGGGCGCCCGCCTGTTGGGACCACTGCGGCGCTGGGAGTTGGGAGTTGGTGC 135
 QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
 Db 136 GAGGCAACTACACTATCGCTGCTGCGAGTTCTTACAAACAGATCAGCAACGTC 195
 QY 41 LeuPhePheIleLeuProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
 Db 196 TTGTTTTCATTTTACCTCCCATCTGCATGCTGCTTCTCCGCCAGTACCAACGCTTC 255
 QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValValGlyIleGlySerValTyr 80
 Db 256 AACACGGCGCATCTACTTAATATGGAGCTCTAGTTGTATGGGGGATGGATCTGTCTAC 315
 QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
 Db 316 TTCCATGCAACGCTGAGTTTCTGGGTGAGATGCTTGTGATGAACTTCCCATCTCTGGGTT 375
 QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
 Db 376 CTGATGTGTGCTTTGGCCATGTGGTTTCCAGGAGGTATTTACCAAAGATCTTTCCGAAT 435
 QY 121 AspArgGlyArgPheLysValValValSerValSerAlaValThrThrCysLeuAla 140
 Db 436 GAC----- 438
 QY 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
 Db 438 ----- 438
 QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
 Db 439 -----AGGTGTGACAAATGTGCGTGTGTGTTAAAGCTGGGCGCTCTTC 477
 QY 181 SerGlyLeuTrpTrpThrLeuAlaIlePheCysTrpIleSerAspArgAlaPheCysGlu 200
 Db 478 CTGGCGCTCTGGTGGACTCTGGCTCTCTTCTGCTGGATCAGGACCAACGCTTCTGTGAG 537
 QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu 220
 Db 538 CTGCTCTCTCTCTTTCACCTTCCCTACCTGCACTGTGTGTGGCATATTCTCATCTGCCIT 597

```

QY 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGlu 240
Db 598 GCITCGPACTGGCTGTGTGCTTCCCTACTTCTTGATGCTGCTCCCTCAGAGATACCTGAG 657
QY 241 GlnGlyProValLeuLysPheTyrProAsnGluLysTyrAlaPheLeuGlyValProTyr 260
Db 658 CAAGGTCCAGTCATCAGATTCTGCCCGCAGCGAGAAATGGCTTTTATTGGTGTCCCTTAT 717
QY 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysLeuThr 275
Db 718 GTGTCCCTTCTGTGTGCCCAAGAAAGTGCCTCAGATCAAGATCAG 762

RESULT 6
AK116177 1429 bp mRNA linear INV 30-NOV-2002
LOCUS Clona intestinalis cDNA, clone:citb013n14, full insert sequence.
DEFINITION AK116177
ACCESSION AK116177.1 GI:23588622
VERSION FLI CDNA.
KEYWORDS Clona intestinalis
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
REFERENCE 1
AUTHORS Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasakura,Y.,
Nakayama,A., Isnikawa,H., Inaba,K. and Satoh,N.
TITLE A cDNA resource from the basal chordate Clona intestinalis
JOURNAL Genesis 33 (4), 153-154 (2002)
MEDLINE 22191024
PUBMED 12203911
REFERENCE 2 (bases 1 to 1429)
AUTHORS Satou,Y. and Satoh,N.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(email:satoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
Fax:81-75-705-1113)
COMMENT Clona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES
source
1. .1429
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb013n14"

ORIGIN
Alignment Scores:
Pred No.: 1,63e-67 Length: 1429
Score: 793.50 Matches: 133
Percent Similarity: 70.83% Conservative: 54
Best Local Similarity: 50.38% Mismatches: 74
Query Match: 52.79% Indels: 3
DB: 3 Gaps: 1

US-10-017-410-4 (1-275) x AK116177 (1-1429)
QY 7 TrpAspGlnLeuGlnAlaGlySerSerGluValAspTyrCysGluAsnTyrThrIle 26
Db 95 TGGAGCGACTTTTACCCCATAGTTCTGAAGTGGACTGGTGTGAGAACAAATTACGCCACA 154
QY 27 ValProAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuPro 46
Db 155 ATACCAATTTATGGCGAATTTGGAAACACTGTAGCAATTCATATTTTATCATTTCCA 214
QY 47 ProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeu 66
Db 215 CCCCTGTAACTACTCTGTTCAGCAATATTCACCAAGTGTGCTCATCATTAACCTT 274
QY 67 IleThrThrLeuLeuValValGlyIleGlySerValTyrPheIlePheThrLeuSer 86
Db 275 GTGTGGGTTTTACTCATATTTTGTGGTGGCTCCGTGTATTTTCACTCAACTTTAAGC 334

```

```

QY 87 PheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTyrValLeuMetCysAlaLeuAla 106
Db 335 CTGTGTTGGCAGTTAATGATGAATAGCAATACTATGGGTGTCTCGTCGATAGACC 394
QY 107 MetTyrPheProArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLys 126
Db 395 ACTTGGCTACCAAGCAATATCTCCCTTCGATTCTCAGATCAGACAGCGCAAAATTTCCAA 454
QY 127 ValValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLysProAlaIle 146
Db 455 GCTATAATAGCTGGTGTGCTGTAGTGAGTACTGTCTCTTGTAACCGGAACCTT 514
QY 147 AsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeu 166
Db 515 AACACGGACTATTGTTTCATGTTTGGGTACCAAGAACTGATTGCTGGTGTAGACTTG 574
QY 167 LysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTyrThr 186
Db 575 AGAAGTTGCGAGTGCCTGGTTCATCGTTCAGATCGGTACAAATGTCACTATATGTTGCTGC 634
QY 187 LeuAlaLeuPheCysTyrIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsn 206
Db 635 GTTGCAGTGTGATGTGTGGTGCAGTGATAGATTTTGTGCAATTTCCCTG-----CAG 685
QY 207 PheProTyrLeuHisCysMetTyrHisIleLeuLeuLeuCysLeuAlaIleTyrLeuGlyCys 226
Db 686 TTCCATATCTACACAGTGCATGCGATATCATGTTGTTGCTTCTTACATGGCGTGT 745
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGluGlnGlyProValIleLys 246
Db 746 GTGTGTTACGCTTACTTCTACGCAACAAACGAGGTCCCGAGCAGTGTCTCGTTTAAGG 805
QY 247 PheTyrProAsnGluLysTyrAlaPheIleGlyValProTyrValSerLeuLeuCysAla 266
Db 806 TTCCTGGCGGCACCATGCTCATATGTTCCGCAATTCGGTTCCTTCCTTCCTTCCTTCCTTCCT 865
QY 267 AsnLysLysSer 270
Db 866 AGCGTCAAGTCT 877

RESULT 7
AF323976 1163 bp mRNA linear INV 01-JAN-2002
LOCUS Drosophila melanogaster brain washing (bwa) mRNA, complete cds.
DEFINITION AF323976
ACCESSION AF323976.1 GI:18028134
VERSION AF323976.1
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1163)
Pascual,A., Boquet,I. and Preat,T.
Direct Submission
Submitted (24-NOV-2000) Institut de Neurobiologie Alfred Fessard,
Avenue de la Terrasse. Bat 32/33, Gif-sur-Yvette 91190, France
FEATURES
source
1. .1163
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="38A"
1. .1163
/gene="bwa"
195..1046
/gene="bwa"
/feature="ceramide; BWA"
/codon_start=1
/product="brain washing"
/protein_id="AAL55991.1"

```

/db_xref="GI:18028135"
/translation="MGGMGGGLLDIYAWAEOQLRPGSSPDWCEGNYLISSNIAEFV
NTFSNLFILPVLIMFKYGRFTPGIHVWLLIVGLSSMYPHATLSLIGQLL
DELAIIWVMAAFLFKPKRYPKFKVNDKRTFSWMLLSAIAATGLSWKPKVNAFV
LMFMSVPTVMVLYTELQVSDQVRLGIBSTTVAVAVFCWINDRIFCEAWSSINFP
YLHGPHHIFIFIAAYTVLVLFAFYVESELQRPQLLKYWPKNEFEFGPIFISIRNPG
KALRNTI"

ORIGIN

Alignment Scores:
Pred. No.: 5,96e-60 Length: 1163
Score: 714.50 Matches: 127
Percent Similarity: 68.87% Conservative: 50
Best Local Similarity: 49.42% Mismatches: 79
Query Match: 47.54% Indels: 1
DB: 3 Gaps: 1

US-10-017-410-4 (1-275) x AF323976 (1-1163)

QY 7 TrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrile 26
Db 243 TGGAGCACTAAGCCGAGCTCGCCGCTCGACTGGTGGCAAGCACTACTTGATT 302
QY 27 ValProAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuPro 46
Db 303 TCATCAACATCCCGAGTTCGGAACACGTTTAGCAACTTCCTGTTCTACTGCGCG 362
QY 47 ProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeu 66
Db 363 CCGCTCTGATAATGCTCTCAAGAGTAGCGGAGCTTGTGACGCCGGAATCCAGTC 422
QY 67 IleTrpThrLeuLeuValValGlyIleGlySerValTyrPheHisPheThrLeuSer 86
Db 423 ATCTGGGTCTGCTCATCTGGTGTGGCTGAGTTCGATGCTATCCATGCCACTTGAGT 482
QY 87 PheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCysAlaLeuAla 106
Db 483 CTGATGGCAGCTGCTGACGAACTGGCCATACTCTGGTCTTTCATGGCGGCTTTTCG 542
QY 107 MetTrpPheProArgArgTyrLeuProLysIlePheAsnAspArgGlyArgPheLys 126
Db 543 CTCTCTATCCGAGCGATCTATCCCAAGTTCGTGAAACAGATCGAAACCTTCAGT 602
QY 127 ValValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLysProAlaIle 146
Db 603 TGGCTCATGCTCTTGTGCGCGATTCGTGCGACGGGCTTCTGTTGGTGAAGCCATTGT 662
QY 147 AsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeu 166
Db 663 AACGCTTGTTCATGTCATGAGTGTGCGACCATGATGTCATGCTTACACAGAGCTG 722
QY 167 LysArgCysAspAsnMetArgValPheLysLeuGlyLysPheSerGlyLeuTrpTrpThr 186
Db 723 CAGAGAGTTAGTGACCAAGGGGTACTACCGCTGGGCATCCGATCGACCGCTCTGGCA 782
QY 187 LeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsn 206
Db 783 GTTGGGTGTTCTGCTGGATCATGATACCGGATCTTCTGGAGCGCTTGTCTCGATCAAT 842
QY 207 PheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCys 226
Db 843 TTTCCATACCTGACGGCTTCTGGACATATTCATTTTATAGCCGCTTACACGGTGTG 902
QY 227 ValCysPheAlaTyrPheAspAlaIleSerGluIleProGluGlnIleProValIleLys 246
Db 903 GTGCTCTTGGCTACTTCTACGTGGAATCGGAGTCCGCCAGCACGCGCTGCTGTAAG 962
QY 247 PheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
Db 963 TACTGGCCGAGAACGAGTTCGAGTTC---GGGATACCCCTTCATTCGATC 1010

RESULT 8
AY071232

LOCUS

DEFINITION Drosophila melanogaster 2318 bp mRNA linear INV 20-DEC-2001
ACCESSION AY071232
VERSION AY071232.1 GI:17945605
KEYWORDS FLI_CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE

AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Drensek,D., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
and Celniker,S.

TITLE

JOURNAL Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA

COMMENT

Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

source
1..2318
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/map="3884-3885"
1..2318
/gene="CG13969"
/note="alignment with genomic scaffold AE003665. gene does
not completely overlap longest ORF"
/db_xref="FLYBASE:FBgn0032842"
404..1255
/gene="CG13969"
/note="Longest ORF"
/codon_start=1
/product="RE26924p"
/protein_id="AAL48854.1"
/db_xref="GI:17945606"
/db_xref="FLYBASE:FBgn0032842"
/translation="MGGMGGGLLDIYAWAWEHLRPGSSPDWCEGNYLISSNIAEFV
NTFSNLFILPVLIMFKYGRFTPGIHVWLLIVGLSSMYPHATLSLIGQLL
DELAIIWVMAAFLFKPKRYPKFKVNDKRTFSWMLLSAIAATGLSWKPKVNAFV
LMFMSVPTVMVLYTELQVSDQVRLGIBSTTVAVAVFCWINDRIFCEAWSSINFP
YLHGPHHIFIFIAAYTVLVLFAFYVESELQRPQLLKYWPKNEFEFGPIFISIRNPG
KALRNTI"

gene

CDS

ORIGIN

Alignment Scores:
Pred. No.: 4,18e-59 Length: 2318
Score: 709.50 Matches: 126
Percent Similarity: 68.48% Conservative: 50
Best Local Similarity: 49.03% Mismatches: 80

Query Match: 47.21% Indels: 1
DB: 3 Gaps: 1

US-10-017-410-4 (1-275) x AV071232 (1-2318)

7 TrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCysGluAspAsnThrThrile 26
452 TGGGAGCACCTAAGACCCGGAAGCTCGCCGCTCGACTGGTGGGAAGCAACTACTTGTATT 511
Qy 27 ValProAlaIleAlaGluPheTyAsnThrThrileSerAsnValLeuPhePheIleLeuPro 46
Db 512 TCGTCCAAACATCGCGAGTTCGTGAACACAGTTTACCACTTCTCTGTTTCATCTACTCGCG 571
Qy 47 ProfileCysMetCysLeuPheAspGluTyAlaThrCysLeuAsnSerAspIleTyLeu 66
Db 572 CCGGCTCTTAATGCTCTTCAAGAGTACGAGCGCTTTGTGAGCCCGGAATCCAGCTC 631
Qy 67 IleTrpThrLeuLeuValValGlyIleGlySerValTyPheHisPheThrLeuSer 86
Db 632 ATCTGGGTGCTGCTCATCGTGTGGCTGAGTTCGATGTTCCATGCCACCTTGTAGT 691
Qy 87 PheLeuGlyGlnMetLeuAspGluLeuAlaValLeuThrValLeuMetCysAlaLeuAla 106
Db 692 CTGATTGGCCAGCTGCTGGACAACTGGCCATCTCTGGGTCTTCAATGGCGGCTTTTCG 751
Qy 107 MetTrpPheProArgArgTyLeuProLysIlePheArgAsnAspArgGlyArgPheLys 126
Db 752 CTCCTTCATCCGAGCGGATATCTCCCAAGTTCGTGAAAAAGATCGCAAAACCTTCAGT 811
Qy 127 ValValValSerValSerAlaValThrCysLeuAlaPheValPheProAlaIle 146
Db 812 TGGCTCATGCTTGTGGCGATTGCTGCGAGCGGCTTGTGCTGGTGAAGCCCATTTGT 871
Qy 147 AsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeu 166
Db 872 AACGCTTTTGTCTCATGTTTCTGATGTTGCGGACCATGTTGTAATGCTCTACACAGAGCTG 931
Qy 167 LysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrpThr 186
Db 932 CAGAGATTAGTACACAGAGGCTTACCCTGCGCATCCGATCCAGACCGCTGCGGCT 991
Qy 187 LeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsn 206
Db 992 GTTGGGTTGTTCTGCTGATCAATGACCGGATCTTCTCGAGGCTGCTGCTCGATCAAT 1051
Qy 207 PheProTyLeuHisCysMetTrpHisIleLeuIleCysLeuAlaAlaTyLeuGlyCys 226
Db 1052 TTTCCATACCTGCGAGGCTTCTGGCACATATTTCAATTTTATAGCGCGCTACACGCTGCTG 1111
Qy 227 ValCysPheAlaTyPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLys 246
Db 1112 GTGCTTTTGGCTACTTCTACGTGAATCGAGCTGCCCCAGCGACACGCGCTGCTGAAG 1171
Qy 247 PheTrpProAsnGluTyTrpAlaPheIleGlyValProTyTrpValSerLeu 263
Db 1172 TACTGGCCAAAGACGAGTTCGAGTTC---GGGATACCTTTCATTTTCGATC 1219

RESULT 9
AX206797
LOCUS
DEFINITION
Sequence 4 from Patent WO0155408.
ACCESSION
AX206797
VERSION
AX206797.1 GI:15394641
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hofmann, K. and Conradt, M.
REFERENCE
AUTHORS
Ceramide
TITLE
Patent: WO 0155408-A 4 02-AUG-2001;
JOURNAL
Memorec Medical Research Cologne GmbH (DE)

FEATURES
source
Location/Qualifiers
1..792
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4,61e-43 Length: 792
Score: 538.50 Matches: 98
Percent Similarity: 60.00% Conservative: 52
Best Local Similarity: 39.20% Mismatches: 99
Query Match: 35.83% Indels: 1
DB: 6 Gaps: 1

US-10-017-410-4 (1-275) x AX206797 (1-792)

Qy 14 SerSerGluValAspTrpCysGluAspAsnThrThrileValProAlaIleAlaGluPhe 33
Db 25 AGCTCCGAGGTGGACTGGTGTGAGAGCAACTTCCAGTACTCGAGCTGGTGGCCGAGTTC 84
Qy 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
Db 85 TACAACACGTTCTCCAAATATCCCTTCTTCACTTTCGGGCGCACATGATGATGCTCTGATG 144
Qy 54 AspGluTyAlaThrCysLeuAsnSerAspIleTyLeuIleTrpThrLeuValVal 73
Db 145 CACCGTATGCCAGAGAGCGCTCCCGCTACATTTACGTTGTCTGGTCTCTTCATGATC 204
Qy 74 ValGlyIleGlySerValTyPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 205 ATAGGCTGTGTTCTCCATGTAATTTCCACATGACGCTCAGCTTCTCTGGGCGAGCTGGAC 264
Qy 94 GluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyr 113
Db 265 GAGATCGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 324
Qy 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeuSer 133
Db 325 TTCCT 384
Qy 134 AlaValThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThr 153
Db 385 GTGGTCAGCACCTTCTGCT 444
Qy 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArg 173
Db 445 ATTGCCCTGACATCTCTCTACATCGTGTGCCAGGAGTACAGGAAGACCCAGCAATAAGGAG 504
Qy 174 ValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAlaLeuPheCysTrpIle 193
Db 505 CTTCGGCACCTGATTGAGTCTCCGTGGTGTATGGGCTGTGTCTCTGACGAGCTGGATC 564
Qy 194 SerAspArgAlaPheCysGluLeuSerSerPheAsnPheProTyLeuHisCysMet 213
Db 565 AGTGACCGCTGCTTTCAGCTTCTGGCAGAGGATTCATTTCTTCTATCTGACAGCATC 624
Qy 214 TrpHisIleLeuIleCysLeuAlaAlaTyLeuGlyCysValCysPheAlaTyTrpPheAsp 233
Db 625 TGGCATGTGCTCATCAGCATCCTCTTATGTCATGGTTCATCCATGTCGCTTGTGGTATC 684
Qy 234 AlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrp 253
Db 685 GCCAACTATGATGCCAGGTGAAACCTCAAGTCCGCTACTGCGCTCGGACAGCTGG 744
Qy 254 AlaPheIleGlyValProTyTrpValSerLeu 263
Db 745 CCC---GTGGGCTGCTCTACGTGGAAATC 771

RESULT 10
AF347024
LOCUS
DEFINITION
Homo sapiens alkaline ceramidase mRNA, complete cds.
795 bp mRNA linear PRI 02-MAR-2002

```
ACCESSION AF347024
VERSION AF347024.1 GI:19070366
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 795)
AUTHORS Mao, C., Xu, R. and Obeid, L.M.
TITLE Molecular cloning and characterization of a novel human alkaline
ceramidase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 795)
AUTHORS Mao, C., Xu, R. and Obeid, L.M.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2001) Medicine, Medical University of South
Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA
FEATURES
source
1..795
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..795
/function="deacylates ceramide to generate sphingosine and
a fatty acid"
/codon_start=1
/product="alkaline ceramidase"
/protein_id="AAL83822.1"
/db_xref="GI:19070366"
MLLMHPYAKRSYIYVWVLFNIGLFSMFHMLTSLFQGLLEIALLGLSGYSI
WMPCRFPLGNRSQIRLVITTTVSTLLSFRLPTVNAALNSIALHLYIVCOE
YRKTNSKRLHLEVLWALVTSIDRLICSFQRIHPFHLISIHVLISITFP
YGMVTMALVDANYEMPGETILKRYWPRDSWVGLPYVEIRGDDKDC"
ORIGIN
Alignment Scores:
Pred. No.: 4,63e-43 Length: 795
Score: 538.50 Matches: 98
Percent Similarity: 60.00% Conservative: 52
Best Local Similarity: 39.20% Mismatches: 99
Query Match: 35.83% Indels: 1
DB: Gaps: 1
US-10-017-410-4 (1-275) x AF347024 (1-795)
QY 14 SerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 25 AGCTCCGAGGTGACTGGTGTGAGACCACTTCCAGTACTCGGAGCTGGTGGCCGAGTTC 84
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
Db 85 TACACACAGTCTCCCAATATCCCTTCTTCATCTTCGGGCACTGATGATGCTCTGATG 144
QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuValVal 73
Db 145 CACCGATGATGCCAGAGCGCTCCCGTCACTTACGTGTGCTGGTCTCTTCATGATC 204
QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 205 ATAGGCTGTCTCCATGATTTTCCATCATGACGCTCAGCTTCTTGGGCCAGCTGGGAC 264
QY 94 GluLeuAlaValLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgTyr 113
Db 265 GAGATGCCATCTGTGCTGCTGGGCACTGGGCTGATGATGATGATGATGATGATGATG 324
QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValValSerValLeuSer 133
Db 325 TTCCCTCTCTTCTGGGGGACAGTCCAGTTCATCGGCTGTCTTCATCACCACCT 384
QY 134 AlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThr 153
Db 385 GTGGTCAGCACCTCTCTGCTCTTCTGGGCGCCACCGTCAACCGCTACGCCCTCAACAGC 444
```

```
QY 154 LeuGlyValProCysThrAlaLeuLeuAlaGluLeuLysArgCysAspAsnMetArg 173
Db 445 ATTGCCCTGCACATTCTTACATCGTGCACAGAGTACAGGAGACCAATAGGAG 504
QY 174 ValPheLysLeuGlyLeuPheSerGlyLeuTyrThrLeuAlaLeuPheCysTyrIle 193
Db 505 CTTGGCAGCCTGATTCAGGTCTCCGTGTTTATGGCTCTTCTGACCACTGATC 564
QY 194 SerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMet 213
Db 565 AGTGACCGCTGCTTGTTCAGCTTCTGCGAGAGATTCACTTCTTATCTGACAGCATC 624
QY 214 TrpHisIleLeuLeuLeuAlaIleTyrLeuGlyCysValCysPheAlaTyrPheAsp 233
Db 625 TGSCATGTGCTCATCAGCATCACCTTCCCTTATGCGATGCTCACCATGGCTTGGTAT 684
QY 234 AlaAlaSerGluLeuProGluGlnGlyProValIleLysPheTyrProAsnGluLysTyr 253
Db 685 GCCAACTATGATGCCAGTGAACCTCAAGTCCGCTACTGCTGCTCGGACAGTTGG 744
QY 254 AlaPheIleGlyValProTyrValSerLeu 263
Db 745 CCC---GTGGGCTGCCCTACTGTGAAATC 771
```

```
RESULT 11
AF347023 822 bp mRNA linear ROD 11-AUG-2003
LOCUS Mus musculus alkaline ceramidase mRNA, complete cds.
DEFINITION AF347023
ACCESSION AF347023
VERSION AF347023.1 GI:19070364
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE 1 (bases 1 to 822)
AUTHORS Mao, C., Xu, R., Szulc, Z.M., Bielawski, J., Becker, K.P., Bielawska, A.,
Galadari, S.H., Hu, W. and Obeid, L.M.
TITLE Cloning and Characterization of a Mouse Endoplasmic Reticulum
Alkaline Ceramidase: AN ENZYME THAT PREFERENTIALLY REGULATES
METABOLISM OF VERY LONG CHAIN CERAMIDES
JOURNAL J. Biol. Chem. 278 (33), 31184-31191 (2003)
PUBMED 12783875
REFERENCE 2 (bases 1 to 822)
AUTHORS Mao, C., Xu, R. and Obeid, L.M.
TITLE Molecular cloning and characterization of a novel murine alkaline
ceramidase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 822)
AUTHORS Mao, C., Xu, R. and Obeid, L.M.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2001) Medicine, Medical University of South
Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA
FEATURES
source
1..822
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..822
/function="deacylates ceramide to generate sphingosine and
a fatty acid"
/codon_start=1
/product="alkaline ceramidase"
/protein_id="AAL83821.1"
/db_xref="GI:19070365"
translation="MHVGTAKMSSIFAYQSSEVDWCSNFQHSSELVAEYNTFSNV
FLFLPGLMFLMHPYAKRTRCFYGVSLFMLIGFSMFHMLTSLFQGLLEIISIL
MLLAGSYVWLPKRFYFGLVITTTTITFTLTFVFTVNAALNSIALH
HLYITVREYKKIRDDRLHIAVSLWMAALTSIDRLICSFQRIHPFHLISIH
HVLISITFPYGIWALVDKAYEMPDKTLKVHWPDRDSWVGLPYVEIQDNKC"
```

ORIGIN

Alignment Scores:
 Pred. No.: 5,27e-41 Length: 822
 Score: 517.50 Matches: 101
 Percent Similarity: 58.33% Conservative: 46
 Best Local Similarity: 40.08% Mismatches: 100
 Query Match: 34.43% Indels: 5
 DB: 10 Gaps: 3

US-10-017-410-4 (1-275) x AF347023 (1-822)

Qy 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
 |||||
 Db 52 AGTTCTGAGCTGGAGTGGTGGAGAGTAATTCAGCAGCTCAGAGTGGTGGCGGAGTTC 111
 |||||
 Qy 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCysLeuPhe 53
 |||||
 Db 112 TACAATACGTTACGAAATGTCTTCTCTCATCTTTGGACCCCTCATGATGTCTCTCATG 171
 |||||
 Qy 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuValVal 73
 |||||
 Db 172 CATCCGTTATGCCAGAGCGTACCGGTGTTCTATGGAGTGTGAGTCTCTTCATGCTC 231
 |||||
 Qy 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
 ::|||
 Db 232 ATAGTCTGTCTCCATGTATTTCCATGACACTCAGCTTCTTGGGACAGCTGCTGGAT 291
 |||||
 Qy 94 GluLeuAlaValLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgTyr 113
 |||||
 Db 292 GAGATCTCCATCTGCTGTGTGGCCAGTGGATACAGTGTGTGGTCCCGGTGTGTAT 351
 |||||
 Qy 114 LeuProIlePheArgAsnAspArgGlyArgPheIysValValSerValLeuSer 133
 |||||
 Db 352 TTTCCTCAGTTCGTCAGGGGACAGGTCTTACTTCAGTCCCTGGTAACTATAACCACT 411
 |||||
 Qy 134 AlaValThrThrCysLeuAlaPheValIysProAlaIleAsnIleSerLeuMetThr 153
 ::|||
 Db 412 ATTATCAGCACCTTCTTGAAGTTCGTCGTAAGCCCAAGTCAATGATATGCTCTCAACAGC 471
 |||||
 Qy 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuIys-----ArgCysAspAsn 171
 ::|||
 Db 472 ATGCCCATCCACATCTCTCATTTGTGGCCAGTGGATACAGTGTGTGGTCCCGGTGTGTAT 531
 |||||
 Qy 172 MetArgValPheIysLeuGlyLeuPheSerGlyLeuTyrTrpThrLeuAlaLeuPheCys 191
 ::|||
 Db 532 CTTCGG-----CATCTGATGTCGGTTCCTGTGCTTATGGCGCGTGCATGACCAAGC 585
 |||||
 Qy 192 TrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHis 211
 ::|||
 Db 586 TGGATCAGTACCGGTACTTTCAGCTTCTTGCAGCTTCTTGCAGCGGATTCACCTTCTACTAGCTGCAC 645
 |||||
 Qy 212 CysMetTrpHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPheAlaTyr 231
 ::|||
 Db 646 AGCATTTGGCCAGCTCCCTCATAGCATCATATTCTTATGGTATGCTGACCATGCGCCCTG 705
 |||||
 Qy 232 PheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGlu 251
 ::|||
 Db 706 GTGGATGCAAGTATGATGATGCCAGATAAACCCCTCAAGTCCACTACTGCGCCCGGAGC 765
 |||||
 Qy 252 LysTrpAlaPheIleGlyValProTyrValSerLeu 263
 |||||
 Db 766 AGCTGGGTGTC---ATCGGGCTACCCCTATGTGGAGATC 798
 |||||

RESULT 12
 AL158206
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP11-363E7 on chromosome 9, complete
 sequence.
 ACCESSION
 AL158206
 VERSION
 AL158206.8
 KEYWORDS
 HTG.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 163542)
 LAIRD, G.
 Direct Submission
 Submitted (25-JUN-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
 requests: clonerequest@sanger.ac.uk
 On Jul 7, 2000 this sequence version replaced gi:8651879.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-363E7 is from the library RPCI-11.2 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see http://bacpac.med.buffalo.edu/
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-363E7. It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-363E7 is at 1 in this sequence. The
 true left end of clone RP11-25202 is at 163443 in this sequence.

FEATURES
 Location/Qualifiers
 source
 1..163542
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-363E7"
 /clone_lib="RPCI-11.2"
 misc_feature
 127756..127796
 /note="match: GSS: Em:AG026881"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,45e-20 Length: 163542
 Score: 335.50 Matches: 93
 Percent Similarity: 54.33% Conservative: 20
 Best Local Similarity: 44.71% Mismatches: 47
 Query Match: 22.32% Indels: 50
 DB: 9 Gaps: 8

US-10-017-410-4 (1-275) x AL158206 (1-163542)
 Qy 80 TyrPheHisPheThrLeuSerPheLeuGlyGln-----MetLeuAspGluLeu 95
 |||||
 Db 17002 TACTTCCATCTTAGCATCTCTGTTGTAATCTCTTGGGCTCAGTCTCTTCTTTGAGGCC 17061
 |||||
 Qy 96 -AlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuPhe 115
 ::|||
 Db 17062 CAAAGGCAATTTATTGTCACATTTCTTCC-----TATTATCC 17097
 |||||
 Qy 115 oLysIlePheArgAsnAspArgGlyArgPheIysValValSerValSerAlaVal 135
 ::|||
 Db 17098 A-----GCAGGACTATTATTAG-----AACATCTTAAACGGGT 17130
 |||||
 Qy 135 lThrThrCysLeuAlaPheValIysProAlaIleAsnAsnIle----- 149

```
Db 17131 CCAAATGCTGAAGATTATCAT-CCTGCTATTCCAGGATGTTGGTAATCTTCATCC 17189
QY 150 -SerLeuMetThrLeuGlyValProCysThrAlaLeuLeuLeuAlaGluLeuLysArgCy 169
Db 17190 TTTCCTTAATAGAGAGGC---CCCTGGGGGTGCAACTACAGTCAGCAAGGTGCGTGGTT 17246
QY 169 sAspAsnMetArg--ValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAl 188
Db 17247 AGACCGGAAGAGAGCAGCTAAACTC----- 17274
QY 188 aleuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPhePr 208
Db 17275 -----AGGCAGCGAGTCTTCATGCTCATCAGGTCTCA-CC 17311
QY 208 oTyrLeuHisCysMetTrpHisIleLeuLeuCysLeuAlaAlaTyrLeuGlyCysValCy 228
Db 17312 TCTGTCTCCCTCGAGGCACATCTTCATCTGCTGCTACTGCGGTGTGTATG 17371
QY 228 sPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTr 248
Db 17372 CTTTGCCTACTTTGATGCTGCTCAGAGATTCTGAGCAAGGCCCTGTCATCAAGTTCTG 17431
QY 248 pProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeuLeuCysAlaAsnLy 268
Db 17432 GCCCAATGAAATGGCCCTTATGTTGTCCTATGTCCTCTGTCGCAACAA 17491
QY 268 sLysSerSerValLysIleThr 275
Db 17492 GAATCATCATGTCAGATCAG 17513
```

```
RESULT 13
AC109451/c
LOCUS
DEFINITION Homo sapiens chromosome 5 clone RP11-170C8, WORKING DRAFT SEQUENCE,
9 unordered pieces.
ACCESSION AC109451
VERSION AC109451.1 GI:18483447
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 219679)
AUTHORS DOE Joint Genome Institute.
DIRECT SUBMISSION
SUBMITTED (04-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 471701
Center clone name: RPCI-11_170C8
-----
Summary Statistics
Consensus quality: 211691 bases at least Q40
Consensus quality: 214895 bases at least Q30
Consensus quality: 216278 bases at least Q20
Estimated insert size: 217000; agarose-fp estimation
Quality coverage: 5.43 in Q20 bases; agarose-fp estimation
Quality coverage: 5.38 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
```

```
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1349: contig of 1349 bp in length
* 1350 1449: gap of unknown length
* 1450 1450: contig of 5852 bp in length
* 7302 7401: gap of unknown length
* 7402 17687: contig of 10286 bp in length
* 17688 17787: gap of unknown length
* 32087 32087: contig of 14300 bp in length
* 32188 32187: gap of unknown length
* 50238 50238: contig of 18051 bp in length
* 50339 50338: gap of unknown length
* 50339 50339: contig of 19356 bp in length
* 69595 69794: gap of unknown length
* 69795 101826: contig of 32032 bp in length
* 101827 101926: gap of unknown length
* 101927 149855: contig of 47929 bp in length
* 149856 149856: gap of unknown length
* 149956 219679: contig of 69724 bp in length.
```

FEATURES

source

```
1. 219679
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-170C8"
/clone_lib="RPCI human BAC library 11"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,07e-20 Length: 219679
Score: 335.50 Matches: 93
Percent Similarity: 54.33% Conservative: 20
Best Local Similarity: 44.71% Mismatches: 47
Query Match: 22.32% Indels: 50
DB: 2 Gaps: 8

US-10-017-410-4 (1-275) x AC109451 (1-219679)
QY 80 TyrPheHisPheThrLeuSerPheLeuGlyGln-----MetLeuAspGluLeu 95
Db 97451 TACTTCCATCTTAGCATCTCTGTGTAATCTCTGGGCTCAGTCTCTTGTGTAGGCC 97392
QY 96 -AlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuPr 115
Db 97391 CAAGGCAITTAITTCACATTTGTTC-----TATTATCC 97356
QY 115 oLysIlePheArGAsnAspArgGlyArgPheLysValValSerValSerAlaVa 135
Db 97355 A-----GCAGGACTATTTAAG-----AACATTCTAAACGGGT 97323
QY 135 lThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIle----- 149
Db 97322 CCCAAATTGCTGAGATTATCAT-CCTGCTATTCCAGGATTTGGTAATCTTCATCC 97264
QY 150 -SerLeuMetThrLeuGlyValProCysThrAlaLeuLeuAlaGluLeuLysArgCy 169
Db 97263 TTTCCTTAATAGAGAGGC---CCCTGGGCTCACTCAGTCAGCAAGGTGCGTGGTT 97207
QY 169 sAspAsnMetArg---ValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAl 188
Db 97206 AGACCGGAAGAGAGCAGCTAAACTC----- 97179
QY 188 aleuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPhePr 208
Db 97178 -----AGGCAGCGAGTCTTCATGCTCATCAGGTCTCA-CC 97142
QY 208 oTyrLeuHisCysMetTrpHisIleLeuLeuCysLeuAlaAlaTyrLeuGlyCysValCy 228
Db 97141 TCTTGTCTCCCTCTGAGGCACATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 97082
```

Qy 228 sPheAlaTyPheAspAlaSerGluIleProGluGlnGlyProValIleLysPheTr 248
 Db 97081 CTTTGGCTACTTTGATGCTGCTCAGAGATTCCTGACAGAGCCCTGTCATCAGTTCG 97022
 Qy 248 pProAsnGluTyPheAlaPheIleGlyValProTyPheValSerIleuLeuCysAlaAsnLy 268
 Db 97021 GCCCAATGAGAAATGGCCCTCATTTGGTGTCCCTATGTCTCCCTCTGTGTGCCAACAA 96962
 Qy 268 slySerSerValIleAlaThr 275
 Db 96961 GAAATCATCATGCAAGATCAGC 96940

RESULT 14
 AC120245 184163 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-259012, WORKING DRAFT SEQUENCE, 5
 DEFINITION unordered pieces.
 AC120245
 VERSION AC120245.4 GI:25008787
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 184163)
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, J., Cesar, H., Center, A.,
 Chacko, J., Chavez, B., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Cartroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,
 Hollins, B., Howells, S., Huix, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokenem, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shivartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 184163)
 Worley, K.C.
 Direct Submission
 Submitted (05-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 184163)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 15, 2002 this sequence version replaced gi:22856351.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GVIT
 Center clone name: CH230-259012
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 175276 bases at least Q40
 Consensus quality: 176285 bases at least Q30
 Consensus quality: 177113 bases at least Q20
 Estimated insert size: 175972; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

 * 1 177885: contig of 177885 bp in length
 * 177886 177985: gap of unknown length
 * 177986 179377: contig of 1392 bp in length
 * 179378 179477: gap of unknown length
 * 179478 180571: contig of 1094 bp in length
 * 180572 180671: gap of unknown length
 * 180672 182236: contig of 1565 bp in length
 * 182237 182336: gap of unknown length
 * 182337 184163: contig of 1827 bp in length.

----- Location/Qualifiers
 1. 184163
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-259012"
 1. 1513
 /note="wgs end extension
 clone_end:Sp6"

----- misc_feature
 misc_feature
 source


```

----- Project Information
Center project name: GAHA
Center clone name: CH230-3E15
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 241576 bases at least Q40
Consensus quality: 243506 bases at least Q30
Consensus quality: 245010 bases at least Q20
Estimated insert size: 253391; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 248465: contig of 248465 bp in length
* 248466 248565: gap of unknown length
* 248566 249734: contig of 1169 bp in length.
FEATURES             Location/Qualifiers
     source            1..249734
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-3E15"
     misc_feature      1..2605
                     /note="wgs_contig"
ORIGIN
Alignment Scores:
Pred No.:           9.67e-19      Length:      249734
Score:             319.00         Matches:     58
Percent Similarity: 75.00%        Conservative: 4
Best Local Similarity: 70.83%     Mismatches:  16
Query Match:       21.22%         Indels:      8
DB:                2              Gaps:         3

US-10-017-410-4 (1-275) x AC097362 (1-249734)
Qy  181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  148948 TCCCGGTGCACTGGACA-----GTGGGGAGGAGGGCGGCCCCAGAG 148889

Qy  201 LeuLeuSerSerPheAsnPhe---ProTyrLeuHisCysMetTrpHisIleLeuIleCys 219
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  148890 AAGCACTGCAGCACCCGCTCACCCCTCCCTCCCTTCCTGCAGG---CACATCCTCATCTGC 148946

Qy  220 LeuAlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIlePro 239
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  148947 CTTGCCGGTACCTGGGCTGGTGTGCTTCGCCTACTTCGATGCTGCTCAGAGATACCC 149006

Qy  240 GluInGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValPro 259
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  149007 GAACAAGGTCCAGTCATCAGATTCTGGCCCAAGTGAAGAAATGGCTTTTATCGTGTCCCC 149066

Qy  260 TyrValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  149067 TATGTGTCCTCTGTGTGTGCCCAAGAAGTGCAGTCAGATCAGC 149114

```

Search completed: September 18, 2004, 06:18:55
Job time : 3427 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 06:19:02 ; Search time 3207 Seconds
(without alignments)
3716.662 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 275
Sequence: 1 MGAHHWDDQLQASSEVDWC.....IGVPVSLLCANKSSVKIT 275

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Pgapop 6.0 , Pgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cpn2.1/USPTO.spool.p/US10017410/runat_15092004_164736_4858/app_query.fasta_1.455
-DB=GenEmbl -OPMT=fastcap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410@CGN.1.1.5600 @runat_15092004_164736_4858 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	136	49.5	2852	9 AK123581	AK123581 Homo sapi
	85	30.9	822	6 AX206799	AX206799 Sequence
c 3	85	30.9	149462	9 AC017081	AC017081 Homo sapi
4	61	22.2	163542	9 AL158206	AL158206 Human DNA
c 5	61	22.2	219679	2 AC109451	AC109451 Homo sapi
6	47	17.1	219679	2 AC109451	AC109451 Homo sapi
7	42	15.3	4078	10 BC059819	BC059819 Mus muscu
8	42	15.3	4174	10 AF282864	AF282864 Mus muscu
9	39	14.2	112713	9 AL391834	AL391834 Human DNA
c 10	36	13.1	358	6 AR419744	AR419744 Sequence
c 11	36	13.1	358	6 BD115297	BD115297 EST and e
12	33	12.0	184163	2 AC120245	AC120245 Rattus no
13	33	12.0	249734	2 AC097362	AC097362 Rattus no
14	25	9.1	90370	10 BX005084	BX005084 Mouse DNA
15	25	9.1	207360	2 BX324222	BX324222 Mus muscu
16	13	4.7	531	11 G76816	G76816 S208P606FG
17	11	4.0	799	8 BT008549	BT008549 Arabidops
18	11	4.0	1193	8 AY090947	AY090947 Arabidops
19	11	4.0	1204	8 BT008652	BT008652 Arabidops
20	11	4.0	1235	8 AB063253	AB063253 Arabidops
21	11	4.0	82891	8 ATT10114	AT021712 Arabidops
c 22	11	4.0	182295	2 BX664738	BX664738 Danio rer
c 23	11	4.0	193226	5 AC144710	AC144710 Danio rer
c 24	11	4.0	199577	8 ATCHRIV57	AL161557 Arabidops
c 25	11	4.0	210673	2 BX769176	BX769176 Danio rer
c 26	9	3.3	906	9 HSA323277	AJ323277 Homo sapi
c 27	9	3.3	12473	9 AF195849S2	AF195844 Homo sapi
c 28	9	3.3	12960	6 AX602161	AX602161 Sequence
c 29	9	3.3	20815	1 AE014271	AE014271 Streptoco
c 30	9	3.3	36711	3 CER0784	Z67756 Caenorhabdi
c 31	9	3.3	42942	7 AB045978	AB045978 Staphyloc
c 32	9	3.3	47289	6 AX059544	AX059544 Sequence
c 33	9	3.3	48012	6 AX702665	AX702665 Sequence
34	9	3.3	48254	6 AX276220	AX276220 Sequence
35	9	3.3	64799	5 AL732411	AL732411 Zebrafish
36	9	3.3	72675	9 BX322573	BX322573 Human DNA
c 37	9	3.3	78153	3 CEY53C10A	AL033536 Caenorhab
c 38	9	3.3	91995	9 AC079914	AC079914 Homo sapi
39	9	3.3	122050	1 SAG766853	AF176685 Streptoco
40	9	3.3	122529	8 TLJ24	AF174763 Arabidops
c 41	9	3.3	126027	8 AC139168	AC139168 Oryza sat
42	9	3.3	126619	9 AC092458	AC092458 Homo sapi
c 43	9	3.3	128000	2 AC004046	AC004046 Homo sapi
c 44	9	3.3	143595	8 AC135209	AC135209 Oryza sat
c 45	9	3.3	144554	9 CNS01DVK	AL135878 Human chr

ALIGNMENTS

RESULT 1

```

AK123581
LOCUS AK123581 2852 bp mRNA linear PRI 09-SEP-2003
DEFINITION Homo sapiens cDNA FLJ741587 fis, clone CTONG2020638.
ACCESSION AK123581
VERSION AK123581.1 GI:34529163
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 2852)
TITLE Isogai, T. and Yamamoto, J.
JOURNAL Direct Submission
COMMENT Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(BE-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..2852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2020638"
/tissue type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 1.08e-135 Length: 2852
Score: 136.00 Matches: 191
Percent Similarity: 98.96% Conservative: 0
Best Local Similarity: 98.96% Mismatches: 1
Query Match: 49.45% Indels: 2
DB: 9 Gaps: 0
US-10-017-410-4 (1-275) x AK123581 (1-2852)
QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103
DB 408 ACCCTTAGTTCCTGGGTCAGATGCTGTGATGAACCTTGACGTCCTTTGGGTCTCGATGTGT 467
QY 104 AlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsnArgGly 123
DB 468 GCTTTGGCCATGTGGTTCCTCCACAGAGTATCTACCAAGATCTTTCGGAATGACCGGGGT 527
QY 124 ArgPheLysValValSerValLeuSerAlaValThrThrCysLeu-AlaPheVally 143
DB 528 AGGTTCAAGTGTGGTGTCTGCTGCTGCGGTTACGACGTGCCCGCC-GGCATTTCGCA 586
QY 143 sProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuI 163
DB 587 GCGTCGCATCAACAAACATCTCTGATGACCCCTGGGAGTTCCTTGACGTGCATGCTCAT 646
QY 163 eAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLe 183

```

```

647 CGCAGAGCTAAAGAGGTGTGACAAACATGGCTGTGTTTAAAGCTGGGCTCTTCTCGGGCT 706
183 uTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSe 203
707 CTGGTGGACCTGGCCCTGTCTCTGTGATCAGTGACCGAGCTTTCTCGAGCTCTGTCTC 766
203 rSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuLeuLeuLeuAlaTyr 223
767 ATCCTTCAACTTCCCTACCTGCACTGTCATGTGGCACATCTCATCTGCTTGTGCTTA 826
223 rLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGluGlnGlyPr 243
827 CTGGGCTGTGTATGCTTGTCTACTTTGATGCTGCTCAGAGATTCTCTGAGCAAGGCC 886
243 oValIleIysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerIe 263
887 TGTTCATCAAGTCTTGGCCCAATGAGAAATGGCCCTTCATTGCTGCTCCCTATGTGCT 946
263 uLeuCysAlaAsnLysLysSerValLysIleThr 275
947 CCTGTGTGCCACAGAAATCATGATCAAGATCAG 983
RESULT 2
AX206799
LOCUS AX206799 822 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 6 from Patent WO0155408.
ACCESSION AX206799
VERSION AX206799.1 GI:15394643
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Hofmann, K. and Conradt, M.
TITLE Ceramidase
JOURNAL Patent: WO 0155408-A 6 02-AUG-2001;
Memorec Medical Molecular Research Cologne Stoffel GmbH (DE)
FEATURES
source
1..822
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.52e-81 Length: 822
Score: 85.00 Matches: 85
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.91% Indels: 0
DB: 6 Gaps: 0
US-10-017-410-4 (1-275) x AX206799 (1-822)
QY 165 GluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrp 184
DB 490 GAGCTAAAGAGGTGTGACAAACATGGCTGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGTG 549
185 TrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSer 204
550 TGGACCTGGCCCTGCTTCTGTGATCAGTGCAGCGAGCTTCTCGAGCTGCTGTCTATCC 609
205 PheAsnPheProTyrLeuHisCysMetTrpHisIleLeuLeuLeuAlaTyrLeu 224
610 TTCAACTTCCCTACCTGCACTGTCATGTGGCACATCTCATCTGCTTGTGCTACCTG 669
225 GlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGluGlnGlyProVal 244
670 GGCTGTGATGCTTTGGCTACTTTGATGTGCTGCTGAGATTCCTGAGCAAGGCCCTGTC 729
245 IleLysPheTrpPro 249

```



```

/rpt_family="(TTG)n"
9685..9973
/rpt_family="Alu"
9974..10197
/rpt_family="L1"
10280..10305
/rpt_family="(TTAG)n"
10320..10341
/rpt_family="AT-rich"
10333..10679
/rpt_family="L1"
10681..10757
/rpt_family="L1"
10758..11039
/rpt_family="Alu"
11040..12011
/rpt_family="L1"
11999..12591
/rpt_family="L1"
12592..12706
/rpt_family="Alu"
12707..12952
/rpt_family="L1"
12966..13140
/rpt_family="L1"
13139..16061
/rpt_family="L1"
16063..17107
/rpt_family="L1"
17108..17161
/rpt_family="MaLR"
17163..17468
/rpt_family="MERL_type"
17471..17675
/rpt_family="MaLR"
17677..17835
/rpt_family="L1"
17909..18144
/rpt_family="L1"
18177..18501
/rpt_family="L1"
18519..20393
/rpt_family="L1"
20490..20906
/rpt_family="L1"
21258..21365
/rpt_family="MIR"
21353..21393
/rpt_family="(T)n"
21366..21684
/rpt_family="Alu"
21543..21566
/rpt_family="AT-rich"
21685..21819
/rpt_family="MIR"
22160..22640
/rpt_family="MER2_type"
22878..23228

```

Alignment Scores:

Pred. No.:	1.89e-79	Length:	149462
Score:	85.00	Matches:	85
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	30.91%	Indels:	0
DB:	9	Gaps:	0

US-10-017-410-4 (1-275) x AC017081 (1-149462)

QY 165 GluLeuLysArgCysAspMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrp 184
 Db 136431 GAGCTAAAGAGGTGTGACAAACATGCGTGTGTTTAAAGCTGGGCGCTTCTCGGCGCTCTGG 136372

QY 185 TrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSer 204
 Db 136371 TGGACCTGGCCCTGTTCTGCTGGAACAGTACCGAGCTTTCGCGAGCTGCTGTCAICC 136312
 QY 205 PheAsnPheProTyrLeuHisCysMetTrpHisIleLeuLeuLeuLeuAlaTyrLeu 224
 Db 136311 TTCAACTTCCTTACCTGACCTGCATGTGGCACATCTCTCATCTGCTTGTCTGCTACCTG 136252
 QY 225 GlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGluGlnGlyProVal 244
 Db 136251 GGCTGTGTATGCTTTGCTTACTTTGATGTGCTGCTCAGAGATTCTCGAGCAAGGCCCTGTC 136192
 QY 245 IleLysPheTrpPro 249
 Db 136191 ATCAAAATTTGGCCC 136177

RESULT 4

AL158206

LOCUS

DEFINITION Human DNA sequence from clone RP11-363E7 on chromosome 9, complete sequence.

ACCESSION

AL158206

VERSION

AL158206.8

KEYWORDS

HTG.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

misc_feature						127756..127796						/note="match: GSS: Em: AG026881"					
ORIGIN																	
Alignment Scores:																	
Pred. No.: 1.28e-53						Length: 163542											
Score: 61.00						Matches: 61											
Percent Similarity: 100.00%						Conservative: 0											
Best Local Similarity: 100.00%						Mismatches: 0											
Query Match: 22.18%						Indels: 0											
DB: 9						Gaps: 0											
US-10-017-410-4 (1-275) x ALI58206 (1-163542)																	
QY						215 HislleLeuilleCysLeuAlalaTyrlEuGllyCyValCysPheAlaTyrPheAspAla						234					
Ddb						17331 CACATCTCATCGCTTGCTGCCACCCTGGGTGTGAAGTTTCCTCCACTTTCATGCT						17390					
QY						235 AlaSerGuilleProGluGlnGlyProValIlleLysPheTrpProAsnGluLysTrpAla						254					
Db						17391 GCCTCAGATTCTGCAGCAAGCCCTGTCATCAAGTTCTGGCCCAATGAGAATATGGGCC						17450					
QY						255 PheIlleGlyValProTyrValSerLeuLeuCysAlaAsnLysLySerserVallylsile						274					
Db						17451 TTCAITGGTGGTCCCTATGTGCTCCCTGCTGTGGCCAAACAAGAAATCATCAGTCAAGATC						17510					
QY						275 Thr 275											
Db						17511 ACG 17513											
RESULT 5																	
AC109451/c																	
LOCUS																	
DEFINITION						Homo sapiens chromosome 5 clone RP11-170C8, WORKING DRAFT SEQUENCE,											
						9 unordered pieces.											
ACCESSION																	
VERSION																	
KEYWORDS						HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.											
SOURCE																	
ORGANISM																	
						Homo sapiens											
						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
						Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
						(bases 1 to 219679)											
						DOE Joint Genome Institute.											
						Sequencing of Human Chromosome 5											
						Unpublished											
						(bases 1 to 219679)											
						DOE Joint Genome Institute.											
						Direct Submission											
						Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint											
						Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA											
						-----Genome Center											
						Center: Joint Genome Institute											
						Center Code: JGI											
						Web site: http://www.jgi.doe.gov											

						Project Information											
						Center Project Name: 471701											
						Center clone name: RPCI-11_170C8											

						Summary Statistics											
						Consensus quality: 211691 bases at least Q40											
						Consensus quality: 214895 bases at least Q30											
						Consensus quality: 216278 bases at least Q20											
						Estimated insert size: 217000; agarose-fp estimation											
						Quality coverage: 5.43 in Q20 bases; sum-of-contigs estimation											
						Quality coverage: 5.38 in Q20 bases; sum-of-contigs estimation.											
						* NOTE: This is a 'working draft' sequence. It currently											
						consists of 9 contigs. The true order of the pieces											
						is not known and their order in this sequence record is											
						arbitrary. Gaps between the contigs are represented as											
						runs of N. But the exact sizes of the gaps are unknown.											

AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	-----Genome Center Center: Joint Genome Institute Center Code: JGI Web site: http://www.jgi.doe.gov

```
/clone_lib="RPCI human BAC library 11"
```

Alignment Scores:	
Pred. No.:	1.82e-38
Score:	47.00
Length:	219679
Percent Similarity:	Matches: 47
Best Local Similarity:	Conservative: 0
Query Match:	Mismatches: 0
DB:	Indels: 0
	Gaps: 0
US-10-017-410-4 (1-275)	x AC109451 (1-219679)

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:69583 IMAGE:6839525"
 /tissue_type="Brain"
 /clone_lib="NIH BMAP_GH0"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1. .4078
 /genes="CRG-L1"
 /db_xref="LocusID:230379"
 76_-765
 /codon_start=1
 /product="CRG-L1 protein"
 /protein_id="AAH59819.1"
 /db_xref="GI:37590520"
 /db_xref="LocusID:230379"
 /translation="MGAPHWMDHLRAGSSVDMCEDNYTIVPAIAFFYNTISNVLFFI
 LPICMCLPROVATCFNSGIVLYLWTLVVVGISGVVFHATLSFLGQMLDELAILVLM
 CALAMFFPRYLKIFRNDRGRPKAVCVLSAITTCGLATKPAINNLSLMILGLPCTA
 LLVAELKRCNVRFKGLFSLGWLTLALFCWISDQAFCELLSSFH
 FPHLCVWHILICLASYLGCVCFAFYDAASEIPEQGVIRFWSEKFAFVGVPVYVSL
 CAHKKSPVKIT"

Alignment Scores: Length: 4078
 Pred. No.: 1.58e-34 Matches: 42
 Score: 42.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x BC059819 (1-4078)

QY 12 AlaGlySerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAala 31
 |||||
 Db 109 GCTGGCAGTTCGGAGTGATGGTGGAGGAGCACTACACTATCGTGCCTGCCATTGCC 168
 |||||

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProPheCysMetCys 51
 |||||
 Db 169 GAGTTCTACACACGATCAGCAGCTCTTTTTCATTTTACCTCCCATCTGCATGTC 228
 |||||

QY 52 LeuPhe 53
 |||||
 Db 229 TTGTTC 234

RESULT 8
 AF282864 4174 bp mRNA linear ROD 12-DEC-2001
 LOCUS Mus musculus cancer related gene-liver 1 mRNA, complete cds.
 DEFINITION
 ACCESSION AF282864
 VERSION AF282864.1 GI:17529683
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4174)
 Graveel,C.R., Jatkoe,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
 Expression profiling and identification of novel genes in
 hepatocellular carcinomas
 Oncogene 20 (21), 2704-2712 (2001)
 21313787
 MEDLINE 11420682
 PUBLISHED
 REFERENCE 2 (bases 1 to 4174)
 Graveel,C.R., Jatkoe,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
 Identification of genes deregulated in murine hepatocellular
 carcinomas using oligonucleotide microarrays and representational
 difference analysis
 JOURNAL
 TITLE
 JOURNAL
 REFERENCE 3 (bases 1 to 4174)
 Farnham,P.J. and Graveel,C.R.

TITLE Direct Submission
 JOURNAL Submitted (27-JUN-2000) Oncology, University of Wisconsin, 1400
 University Avenue, Madison, WI 53706, USA
 FEATURES
 source
 1. .4174
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="hepatocellular carcinoma"
 35. .862
 /note="CRG-L1; putative"
 /codon_start=1
 /product="cancer related gene-liver 1"
 /protein_id="AAL40408.1"
 /db_xref="GI:17529684"
 /translation="MGAPHWMDHLRAGSSVDMCEDNYTIVPAIAFFYNTISNVLFFI
 LPICMCLPROVATCFNSGIVLYLWTLVVVGISGVVFHATLSFLGQMLDELAILVLM
 CALAMFFPRYLKIFRNDRGRPKAVCVLSAITTCGLATKPAINNLSLMILGLPCTA
 LLVAELKRCNVRFKGLFSLGWLTLALFCWISDQAFCELLSSFHFPYHLCVWHIL
 CLASYLGCVCFAFYDAASEIPEQGVIRFWSEKFAFVGVPVYVSLICAHKKSFKIT"

CDS
 35. .862
 /note="CRG-L1; putative"
 /codon_start=1
 /product="cancer related gene-liver 1"
 /protein_id="AAL40408.1"
 /db_xref="GI:17529684"

ORIGIN
 Alignment Scores: Length: 4174
 Pred. No.: 1.61e-34 Matches: 42
 Score: 42.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.27% Indels: 0
 DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x AF282864 (1-4174)

QY 12 AlaGlySerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAala 31
 |||||
 Db 68 GCTGGCAGTTCGGAGTGATGGTGGAGGAGCACTACACTATCGTGCCTGCCATTGCC 127
 |||||

QY 32 GluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProPheCysMetCys 51
 |||||
 Db 128 GAGTTCTACACACGATCAGCAGCTCTTTTTCATTTTACCTCCCATCTGCATGTC 187
 |||||

QY 52 LeuPhe 53
 |||||
 Db 188 TTGTTC 193

RESULT 9
 AL391834 112713 bp DNA linear PRI 27-FEB-2001
 LOCUS Human DNA sequence from clone RP11-513M16 on chromosome 9, complete
 DEFINITION
 ACCESSION AL391834
 VERSION AL391834.8 GI:13169552
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 112713)
 Laird,G.
 Direct Submission
 Submitted (27-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Feb 28, 2001 this sequence version replaced gi:13160304.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 corresponding to the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 together with a note of the overlapping clone name. Note that the
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>. RP11-513M16 is from the library RP01-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-513M16. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-513M16 is at 1 in this sequence. The true left end of clone RP11-363E7 is at 112614 in this sequence. The true right end of clone RP11-146N23 is at 31926 in this sequence.

FEATURES

	Location/Qualifiers	
source	1..112713	repeat_region
	/organism="Homo sapiens"	repeat_region
	/mol_type="genomic DNA"	repeat_region
	/db_xref="taxon:9606"	repeat_region
	/chromosome="9"	repeat_region
	/clone="RP11-513M16"	repeat_region
	/clone_lib="RP01-11.2"	repeat_region
repeat_region	1..107	repeat_region
repeat_region	/note="MER33 repeat: matches 1..109 of consensus"	repeat_region
repeat_region	115..151	repeat_region
repeat_region	/note="L1MC5 repeat: matches 7292..7330 of consensus"	repeat_region
repeat_region	311..345	repeat_region
repeat_region	/note="L1MC5 repeat: matches 7193..7227 of consensus"	repeat_region
repeat_region	501..971	repeat_region
repeat_region	/note="L2 repeat: matches 1791..2267 of consensus"	repeat_region
repeat_region	936..1055	repeat_region
repeat_region	/note="L2 repeat: matches 1093..1151 of consensus"	repeat_region
repeat_region	1056..1343	repeat_region
repeat_region	/note="AluX repeat: matches 1..301 of consensus"	repeat_region
repeat_region	1344..1972	repeat_region
repeat_region	/note="L2 repeat: matches 472..1093 of consensus"	repeat_region
repeat_region	1973..2244	repeat_region
repeat_region	/note="AluY repeat: matches 1..301 of consensus"	repeat_region
repeat_region	2245..2515	repeat_region
repeat_region	/note="L2 repeat: matches 205..472 of consensus"	repeat_region
repeat_region	2516..2643	repeat_region
repeat_region	/note="AluJo/FLAM repeat: matches 1..127 of consensus"	repeat_region
repeat_region	2646..2938	repeat_region
repeat_region	/note="AluX repeat: matches 1..298 of consensus"	repeat_region
repeat_region	2939..3101	repeat_region
repeat_region	/note="L2 repeat: matches 11..205 of consensus"	repeat_region
repeat_region	3376..3715	repeat_region
repeat_region	/note="MER46B repeat: matches 1..235 of consensus"	repeat_region
repeat_region	4229..4509	repeat_region
repeat_region	/note="AluJo repeat: matches 1..284 of consensus"	repeat_region
repeat_region	6954..7122	repeat_region
repeat_region	/note="L1ME repeat: matches 5620..5799 of consensus"	repeat_region
repeat_region	7825..8113	repeat_region
repeat_region	/note="AluSq repeat: matches 1..290 of consensus"	repeat_region
repeat_region	8114..8209	repeat_region
repeat_region	/note="4 copies 24 mer 83% conserved"	repeat_region
repeat_region	8115..8200	repeat_region
repeat_region	/note="43 copies 2 mer tc 73% conserved"	repeat_region
repeat_region	8260..8548	repeat_region
repeat_region	/note="AluSP repeat: matches 1..289 of consensus"	repeat_region
repeat_region	8549..9285	repeat_region
repeat_region	/note="L1MB8 repeat: matches 5402..6169 of consensus"	repeat_region
repeat_region	9827..10130	repeat_region
repeat_region	/note="AluSc repeat: matches 1..303 of consensus"	repeat_region
repeat_region	10261..10570	repeat_region
repeat_region	/note="AluSq repeat: matches 1..304 of consensus"	repeat_region
repeat_region	10701..10993	repeat_region
repeat_region	/note="AluSc repeat: matches 1..293 of consensus"	repeat_region
repeat_region	11835..12154	repeat_region
repeat_region	/note="AluX repeat: matches 1..311 of consensus"	repeat_region
repeat_region	12158..12230	repeat_region
repeat_region	/note="AluJ/FRAM repeat: matches 229..295 of consensus"	repeat_region
repeat_region	12264..12391	repeat_region
repeat_region	/note="FLAM C repeat: matches 2..133 of consensus"	repeat_region
repeat_region	12547..12808	repeat_region
repeat_region	/note="AluJo repeat: matches 1..261 of consensus"	repeat_region
repeat_region	12861..13043	repeat_region
repeat_region	/note="MER20 repeat: matches 28..212 of consensus"	repeat_region
repeat_region	13054..13155	repeat_region
repeat_region	/note="MIR repeat: matches 145..250 of consensus"	repeat_region
repeat_region	13161..13258	repeat_region
repeat_region	/note="MIR repeat: matches 115..214 of consensus"	repeat_region
repeat_region	13534..13843	repeat_region
repeat_region	/note="AluSq repeat: matches 1..306 of consensus"	repeat_region
repeat_region	14011..14313	repeat_region
repeat_region	/note="AluX repeat: matches 1..308 of consensus"	repeat_region
repeat_region	14735..15655	repeat_region
repeat_region	/note="L1MA8 repeat: matches 5354..6280 of consensus"	repeat_region
repeat_region	15656..15723	repeat_region
repeat_region	/note="34 copies 2 mer at 72% conserved"	repeat_region
repeat_region	18498..18802	repeat_region
repeat_region	/note="L1MC4 repeat: matches 7639..7970 of consensus"	repeat_region
repeat_region	18829..19693	repeat_region
repeat_region	/note="L1MC4 repeat: matches 6644..7556 of consensus"	repeat_region
repeat_region	19715..19934	repeat_region
repeat_region	/note="L1MC4 repeat: matches 6223..6449 of consensus"	repeat_region
repeat_region	20794..21102	repeat_region
repeat_region	/note="AluJb repeat: matches 1..312 of consensus"	repeat_region
repeat_region	21336..21640	repeat_region
repeat_region	/note="AluSq repeat: matches 1..303 of consensus"	repeat_region
repeat_region	23785..23843	repeat_region
repeat_region	/note="MLT1A1 repeat: matches 291..352 of consensus"	repeat_region
repeat_region	23844..23988	repeat_region
repeat_region	/note="FRAM repeat: matches 1..132 of consensus"	repeat_region
repeat_region	23989..24294	repeat_region
repeat_region	/note="AluJb repeat: matches 2..239 of consensus"	repeat_region
repeat_region	24295..24329	repeat_region
repeat_region	/note="AluJo/FRAM repeat: matches 268..304 of consensus"	repeat_region
repeat_region	24427..24611	repeat_region
repeat_region	/note="MLT1A1 repeat: matches 1..198 of consensus"	repeat_region
repeat_region	24780..24886	repeat_region
repeat_region	/note="L2 repeat: matches 2596..2710 of consensus"	repeat_region
repeat_region	26667..26923	repeat_region
repeat_region	/note="AluX repeat: matches 1..267 of consensus"	repeat_region
repeat_region	27639..27810	repeat_region
repeat_region	/note="MIR repeat: matches 19..193 of consensus"	repeat_region
repeat_region	28304..28591	repeat_region
repeat_region	/note="AluJo repeat: matches 1..298 of consensus"	repeat_region
repeat_region	28632..28896	repeat_region
repeat_region	/note="AluJb repeat: matches 14..282 of consensus"	repeat_region
repeat_region	28974..29101	repeat_region
repeat_region	/note="L2 repeat: matches 2634..2750 of consensus"	repeat_region
repeat_region	30422..30555	repeat_region
repeat_region	/note="FLAM C repeat: matches 1..132 of consensus"	repeat_region
repeat_region	30740..30920	repeat_region
repeat_region	/note="MER5A repeat: matches 18..188 of consensus"	repeat_region
repeat_region	31021..31327	repeat_region
repeat_region	/note="AluX repeat: matches 3..310 of consensus"	repeat_region
repeat_region	31409..31538	repeat_region
repeat_region	/note="L1MC4 repeat: matches 7733..7866 of consensus"	repeat_region
repeat_region	32346..32575	repeat_region
repeat_region	/note="AluJo repeat: matches 1..229 of consensus"	repeat_region
repeat_region	32841..33137	repeat_region
repeat_region	/note="AluX repeat: matches 2..299 of consensus"	repeat_region
repeat_region	33144..33391	repeat_region
repeat_region	/note="AluX repeat: matches 1..245 of consensus"	repeat_region
repeat_region	33872..33978	repeat_region

/note="L1MC/D repeat: matches 5467. .5582 of consensus"
34043. .34338
/note="AluSq repeat: matches 1. .305 of consensus"
34386. .34702
/note="AluY repeat: matches 1. .311 of consensus"
34785. .35090
/note="L1M4 repeat: matches 4840. .5159 of consensus"
35268. .35464
/note="L1M4 repeat: matches 4413. .4621 of consensus"
35800. .35944
/note="L1ME2 repeat: matches 5987. .6145 of consensus"
36282. .36315
/note="L17 copies 2 mer ga 97% conserved"
36980. .37089
/note="L2 repeat: matches 2583. .2696 of consensus"
37777. .38750
/note="L1MC5 repeat: matches 6973. .7941 of consensus"
38751. .38868
/note="FLAM A repeat: matches 5. .121 of consensus"
38869. .39168
/note="L1MC5 repeat: matches 6673. .6973 of consensus"
40016. .40401
/note="MLT1J repeat: matches 107. .512 of consensus"
40402. .40655
/note="AluSq repeat: matches 1. .265 of consensus"
40656. .40690
/note="MLT1J repeat: matches 75. .107 of consensus"

Alignment Scores:
Pred. No.: 4.16e-30 Length: 112713
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.18% Indels: 0
DB: 9 Gaps: 0

US-10-017-410-4 (1-275) x AL391834 (1-112713)

Qy 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103
Db 104219 ACCCTAGTTCTTGGTGCAGATGCTTGATGAATTCGAGTCCTTGGGTTCGTATGTT 104278

Qy 104 AlaLeuAlaMetTrpPheProArgArgTrpLeuProLysIlePheArgAsnAspArg 122
Db 104279 GCTTTGGCCATGTGGTTCCTCCCGAAGGATATCTACCAAGATCTTTCGAGATGACCG 104335

RESULT 10
AR419744/c 358 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 11241 from patent US 6639063.
ACCESSION AR419744
VERSION AR419744.1 GI:40174854
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 358)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 11241 28-OCT-2003;
FEATURES
source
1. .358
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 5.88e-29 Length: 358
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.09% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x AR419744 (1-358)

Qy 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerGluValAspTrpCys 20
Db 111 ATGGGGCGCCCGCAGTGGTGGGACAGCTGCAGGCTGGTAGCTCGAGGTGGACTGGTGC 52

Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThr 36
Db 51 GAGGACAACTACACCATCTGCTGCTATCGCGAGTTCCTACACACG 4

RESULT 11
BD115297/c 358 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION EST and encoded human protein.
ACCESSION BD115297
VERSION BD115297.1 GI:23210201
KEYWORDS JP 2002010789-A/7374.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 358)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 7374 15-JAN-2002;
COMMENT GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/7374
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1. .358
/organism="Homo sapiens (human)".
FEATURES
source
1. .358
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 5.88e-29 Length: 358
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.09% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-4 (1-275) x BD115297 (1-358)

Qy 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerGluValAspTrpCys 20
Db 111 ATGGGGCGCCCGCAGTGGTGGGACAGCTGCAGGCTGGTAGCTCGAGGTGGACTGGTGC 52

Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThr 36
Db 51 GAGGACAACTACACCATCTGCTGCTATCGCGAGTTCCTACACACG 4

RESULT 12
AC120245 184163 bp DNA linear HTG 15-NOV-2002
LOCUS
DEFINITION Rattus norvegicus clone CH230-259012, WORKING DRAFT SEQUENCE, 5
unordered pieces.
ACCESSION AC120245

VERSION
KEYWORDS
SOURCE
ORGANISMAC120245.4 GI-25008787
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)

REFERENCE

AUTHORS

1 (bases 1 to 184163)
Muzny, D., Marle, Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angilano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Z., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Coyle, M., Cree, A., D'Souza, L., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denon, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbini, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loubege, H., Lozado, R.J., Lu, X., Ma, J., Mangum, B., Mapua, P., Martin, K., Mahmoud, M., Malloy, K., Mangum, A., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedeme, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, X., Zhao, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Center project name: GVTT
Center clone name: CH230-259012
Assembly program: Phrap; version 0.990329
Consensus quality: 175276 bases at least Q40
Consensus quality: 176285 bases at least Q30
Consensus quality: 177113 bases at least Q20
Estimated insert size: 175972; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 177885: contig of 177885 bp in length
177886 177985: gap of unknown length
177986 179377: contig of 1392 bp in length
179378 179477: gap of unknown length
179478 180571: contig of 1094 bp in length
180571 180672: gap of unknown length
180672 182336: contig of 1565 bp in length
182336 182337: gap of unknown length
182337 184163: contig of 1827 bp in length.

FEATURES

source

1..184163
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-259012"

misc_feature

1..1513
/note="wgs_end extension
clone_end:Sp6"
complement(4969..5842)
/note="clone boundary
clone_end:Sp6
site:
end sequence:BZ169793"
40368..40445
/note="clone boundary
clone_end:T7
site:
end sequence:BZ169790"

misc_feature

171006..172144
/note="wgs_end extension
clone_end:T7"
173214..174400
/note="wgs_end extension
clone_end:T7"
175979..177885
/note="wgs_end extension"

misc_feature

171006..172144
/note="wgs_end extension
clone_end:T7"
173214..174400
/note="wgs_end extension
clone_end:T7"
175979..177885
/note="wgs_end extension"

misc_feature

171006..172144
/note="wgs_end extension
clone_end:T7"
173214..174400
/note="wgs_end extension
clone_end:T7"
175979..177885
/note="wgs_end extension"

misc_feature

171006..172144
/note="wgs_end extension
clone_end:T7"
173214..174400
/note="wgs_end extension
clone_end:T7"
175979..177885
/note="wgs_end extension"

misc_feature

171006..172144
/note="wgs_end extension
clone_end:T7"
173214..174400
/note="wgs_end extension
clone_end:T7"
175979..177885
/note="wgs_end extension"

The sequence in this assembly is a combination of BAC based reads

clone_end:T7"

ORIGIN

Alignment Scores:

Pred. No.: 1.76e-23 Length: 184163
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.00% Indels: 0
 DB: 2 Gaps: 0

US-10-017-410-4 (1-275) x AC120245 (1-184163)

QY 173 ArgValPheIysLeuGlyLeuPheSerGlyLeuTrrPrrThrLeuAlaLeuPheCysTrrP 192

Db 7765 CGTGTGTTAAGCTGGGCCCTTCTCTCGGCTTGGGGACTCTGGCTCTCTCTGCTGG 7824

QY 193 IleSerAspArgAlaPheCysGluLeuLeuSerSerPhe 205

Db 7825 ATCAGTGACCGAGCCTTCTGTGAGCTGTCTCTCTCTT 7863

RESULT 13

AC097362

LOCUS

Rattus norvegicus clone CH230-3E15, WORKING DRAFT SEQUENCE, 2 linear HTG 09-MAY-2003

AC097362

AC097362

HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 249734)
 Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,X., Chen,Z.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Mitosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwackelmeoh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Woodden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

TITLE

Direct Submission

Unpublished

2 (bases 1 to 249734)

Worley,K.C.

Direct Submission

Submitted (17-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 249734)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 9, 2003 this sequence version replaced gi:24956965.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GAHA

Center clone name: CH230-3E15

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 241576 bases at least Q40

Consensus quality: 243506 bases at least Q30

Consensus quality: 245010 bases at least Q20

Estimated insert size: 253391; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 248465: contig of 248465 bp in length

* 248466 248565: gap of unknown length

* 248566 249734: contig of 1169 bp in length.

Location/Qualifiers

1. 249734

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

FEATURES

source

```
misc_feature
1. .2605
/clone="CH230-3E15"
/notes="wgs_contig"

ORIGIN
Alignment Scores:
Pred. No.: 2,26e-23 Length: 249734
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.00% Indels: 0
DB: 2 Gaps: 0

US-10-017-410-4 (1-275) x AC097362 (1-249734)

QY 173 ArgValPheLysLeuGlyLeuPheSerGlyLeuTriPThrLeuAlaLeuPheCysTyr 192
Db 143092 CGTGTGTTTAAGTCGGCCCTTCTCTGCGCTTGTGGCTTGTGGCTCTGCTCTCTCTCTG 143151

QY 193 IleSerAspArgAlaPheCysGluLeuLeuSerSerPhe 205
Db 143152 ATCAGTGACGAGCCCTCTGTGAGCTGCTGTCTCTCTTT 143190

RESULT 14
BX005084
LOCUS
DEFINITION
Mouse DNA sequence from clone RP24-468M3 on chromosome 4, complete
sequence.
ACCESSION
BX005084
VERSION
BX005084.16 GI:37936459
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 90370)
Barlow,K.
Direct Submission
Submitted (22-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 23, 2003 this sequence version replaced gi:37936442.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP24-468M3 is
from a Male (C57BL/6J) mouse BAC Library VECTOR: pTARBAC1.
Location/Qualifiers

source
1. .90370
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP24-468M3"
/clone_lib="RPCT-24"

ORIGIN
Alignment Scores:
Pred. No.: 3,86e-15 Length: 90370
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.09% Indels: 0
DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x BX005084 (1-90370)

QY 12 AlaGlySerSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAla 31
Db 35175 GCTGGCAGTTCGAGGTGGATTGGTGGAGGACACTACACTATCGTGGCTGCATTGCC 35234

QY 32 GluPheTyrAsnThr 36
Db 35235 GAGTCTTACACACG 35249

RESULT 15
BX324222
LOCUS
DEFINITION
Mus musculus chromosome 4 clone RP23-12709, 4 unordered pieces.
ACCESSION
BX324222
VERSION
BX324222.8 GI:31620741
KEYWORDS
HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207360)
Harrison,E.
Direct Submission
Submitted (08-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 11, 2003 this sequence version replaced gi:3159318.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
Project Information
Center project name: BM12709
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 99% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 206005 bases at least Q40
Consensus quality: 206174 bases at least Q30
Consensus quality: 206291 bases at least Q20
Insert size: 207060; sum-of-contigs
Insert size: 20738; 4.4% error; agarose-fp
Quality coverage: 14.24x in Q20 bases; sum-of-contigs Quality
coverage: 14.27x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
```

Mon Sep 20 11:04:18 2004

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 17663: contig of 17663 bp in length
* 17664 17763: gap of 100 bp
* 17764 61868: contig of 44105 bp in length
* 61869 61969: gap of 100 bp
* 61969 83921: contig of 21953 bp in length
* 83922 84021: gap of 100 bp
* 84022 207360: contig of 123339 bp in length.

FEATURES

source
1..207360
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-12709"
/clone_lib="RPCI-23"
1..17663
/note="assembly fragment:03369"
fragment_chain:1
clone_end:SP6
vector_side:left
17764..61868
/note="assembly fragment:01668"
fragment_chain:1
61969..83921
/note="assembly fragment:04768"
fragment_chain:2
84022..207360
/note="assembly fragment:04990"
fragment_chain:2
clone_end:T7
vector_side:right"

ORIGIN

Alignment Scores:
Pred. No.: 7.7e-15 Length: 207360
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.09% Indels: 0
DB: 2 Gaps: 0

US-10-017-410-4 (1-275) x BX324222 (1-207360)

Qy 98 LeuTnpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysile 117
Db 11407 CTGTGGGTCTGTGATGTGTGCTTTGGCCATGTGGTTCCAGGAGGTATTACCAAGATC 11466
Qy 118 PheArgAsnAspArg 122
Db 11467 TTTCGAATGACAGG 11481

Search completed: September 18, 2004, 08:11:08
Job time : 3519 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 05:13:04 : Search time 86 Seconds
(without alignments)
1774.552 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 1503
Sequence: 1 MGAPHHWDLQAGSSEVDWC.....IGVPYVLLCANKSSVKIT 275

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/p/US10017410/runat_15092004_164710_4725/app.query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410.gCgn 1.1.105 @runat_15092004_164710_4725 -NCFU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	213	14.2	358	US-09-621-976-11241	Sequence 11241, A
C 2	117	7.8	6755	US-08-931-999-4	Sequence 4, Appli
C 3	114	7.6	246240	US-08-724-394A-20	Sequence 20, Appl
C 4	114	7.6	246240	US-08-724-394A-21	Sequence 21, Appl
C 5	114	7.6	246240	US-08-724-394A-22	Sequence 22, Appl
C 6	104.5	7.0	1845	US-09-540-236-938	Sequence 938, Appl
C 7	102.5	6.8	269223	US-09-596-002-41	Sequence 41, Appl
C 8	100.5	6.7	1257	US-09-252-991A-892	Sequence 892, Appl
C 9	100.5	6.7	2859	US-09-252-991A-944	Sequence 944, Appl
C 10	100	6.7	1575	US-09-252-991A-1139	Sequence 1139, Appl
C 11	97	6.5	1260	US-09-252-991A-15016	Sequence 15016, A
C 12	97	6.5	1605	US-09-252-991A-15148	Sequence 15148, A

C 13	97	6.5	1950	4	US-09-252-991A-14616	Sequence 14616, A
C 14	97	6.5	2082	4	US-09-252-991A-2596	Sequence 2596, Ap
C 15	97	6.5	2388	4	US-09-252-991A-2514	Sequence 2514, Ap
C 16	96	6.4	2304	4	US-09-540-236-426	Sequence 426, App
C 17	95.5	6.4	1297	3	US-09-083-521-4	Sequence 4, Appli
C 18	95.5	6.4	1350	4	US-09-328-352-3355	Sequence 3355, Ap
C 19	93.5	6.2	92407	4	US-09-596-002-36	Sequence 36, Appl
C 20	93	6.2	774	4	US-09-252-991A-3254	Sequence 3254, Ap
C 21	93	6.2	1794	4	US-09-328-352-577	Sequence 577, App
C 22	92.5	6.2	1079	1	US-08-348-792-7	Sequence 7, Appli
C 23	92.5	6.2	1079	2	US-08-462-738-7	Sequence 7, Appli
C 24	92.5	6.2	1079	4	US-09-199-955-7	Sequence 7, Appli
C 25	92.5	6.2	1079	4	US-08-880-875-7	Sequence 1101, Ap
C 26	92.5	6.2	1200	4	US-09-489-039A-1101	Sequence 1101, Ap
C 27	92.5	6.2	1338	4	US-09-328-352-2543	Sequence 2543, Ap
C 28	91	6.1	1083	4	US-09-328-352-3753	Sequence 3753, Ap
C 29	91	6.1	1116	4	US-09-252-991A-2320	Sequence 2320, Ap
C 30	89.5	6.0	4792	4	US-08-961-527-156	Sequence 156, App
C 31	89.5	6.0	20674	4	US-09-641-638-651	Sequence 651, App
C 32	89	5.9	1671	4	US-09-221-017B-625	Sequence 625, App
C 33	89	5.9	1668	4	US-09-107-532A-2768	Sequence 2768, Ap
C 34	88.5	5.9	1459	2	US-08-970-133-2	Sequence 2, Appli
C 35	88.5	5.9	1459	4	US-09-294-545-2	Sequence 2, Appli
C 36	88.5	5.9	3748	2	US-08-958-240-1	Sequence 715, App
C 37	88	5.9	838	4	US-09-221-017B-715	Sequence 391, Ap
C 38	88	5.9	933	4	US-09-328-352-3991	Sequence 4, Appli
C 39	88	5.9	1189	1	US-07-781-034-4	Sequence 4, Appli
C 40	88	5.9	1189	5	PCT-US92-08328-4	Sequence 693, App
C 41	88	5.9	3552	4	US-09-134-001C-693	Sequence 3627, Ap
C 42	87.5	5.8	960	4	US-09-328-352-3627	Sequence 1, Appli
C 43	87.5	5.8	1716	1	US-08-348-792-1	Sequence 1, Appli
C 44	87.5	5.8	1716	2	US-08-462-738-1	Sequence 1, Appli
C 45	87.5	5.8	1716	4	US-09-199-955-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-621-976-11241/c
; Sequence 11241, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 11241
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-11241

Alignment Scores:
Pred. No.: 1,618-15
Score: 213.00
Percent Similarity: 100.00%
Best Local Similarity: 97.30%
Query Match: 14.17%
DB: 4
Length: 358
Matches: 36
Conservative: 1
Mismatches: 0
Indels: 0
Gaps: 0

US-10-017-410-4 (1-275) x US-09-621-976-11241 (1-358)

Qy	1	MetGlyAlaProHisTriPAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys	20
Db	111	ATGGGGCGCCCGACCTGGTGGGACCGAGCTGGTAGCTCGAGGTGACTGGTGC	52
Qy	21	GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIle	37

Qy	13	GlySerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGlu	32
Db	16465	GGGAAATCGCTCACTCATTTTAAACAGGACGATGGTTGTAGCTGTTTTTAAAAAGTAT	16524
Qy	33	PheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCysLeu	52
Db	16525	TTTATAACCTTTAAAGAAATGTCGTTTGTCTTCCTTTAATCAATA---TTGTGCTTA	16581
Qy	53	PheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIle---TrrThrLeuLeuVa	72
Db	16582	TGAGACCCATTC-----ATCTATTATATATTGGGGCTGTGTTAAT	16620
Qy	72	IValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetIle	92
Db	16621	TCCTATTGCTATACAAATAGACCGTGGTTATTACTCATTCCTTTTTTAAACAGCAACT	16680
Qy	92	uAsnGluLeuAla-----ValLeuTrpVa	100

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-21

```

Alignment Scores:

```

Pred. No.: 3,77 Length: 246240
Score: 114.00 Matches: 74
Percent Similarity: 33.44% Conservative: 29
Best Local Similarity: 24.03% Mismatches: 95
Query Match: 7.58% Indels: 110
DB: 2 Gaps: 14

```

US-10-017-410-4 (1-275) x US-08-724-394A-21 (1-246240)

```

QY 13 GlySerSerGluValAlaAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGlu 32
Db 16465 GGGAAATCGCTCACTGATTTTAAACAGGACAGATGGTGTAGCTGTTTAAAAAGTAT 16524
QY 33 PheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeu 52
Db 16525 TTTATAAACCTTTAAAGAAATGTCGTTTGTCTCTTTAATCAATA---TTTGCTTA 16581
QY 53 PheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIle-TripThrLeuValVa 72
Db 16582 TGAGACCCCAATC-----ATCTATTATATATTGGGTCGTGTAAT 16620
QY 72 lValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetIle 92
Db 16621 TCTCATTTGCTATACAAATAGACCGTGTATTATTACTCATTTTAAACAGACAACT 16680
QY 92 uAspGluLeuAla-----ValIleTyrVa 100
Db 16681 TCTGTTCAGGCTATTACAGATAGTCTCCTATGACATTTGTTGATCATGTTTAAAGAT 16740
QY 100 lLeuMetCysAla-lLeuAlaMetTrpPheProArg-----A 112
Db 16741 GATATGTCATCTTCAAGTAATAAATGTTTGTAGTCACTGGAATTTGACATTTAGCTTTAGG 16800
QY 112 rGlyLeuProIlyllePheArgAsnAspArgGlyArgPheIlyValValValSerValL 132
Db 16801 AGATATTGCGAGCAGCTTT-----CCAAAGTTTCATCTCGTAACGTATGAAA 16848
QY 132 eSerAlaValThrThrCysLeuAla-----PheValIlyProAlaIleAsnAsnI 149
Db 16849 GTTCAGTTGCTCCACATGCTGGCAACATTTTGTGTTTCCATCTCCTTTCAITTTATCCA 16908
QY 149 lSer-----LeuM 152
Db 16909 TTAGTTAATTTATAGTGGTATTTTCATTTGATTTTTCATTTCTGTAATGACTAA 16968
QY 152 etThrLeuGlyVal-----P 157

```

```

Db 16969 TGAAGTTGAACATCTTTTATATGTTAAGGAGCCACTTACATTTTCTTTTATGAAGTC 17028
QY 157 roCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysL 177
Db 17029 CCTGTTCAAGT-----CATGGCCCATTTTATAGT 17058
QY 177 euGlyLeuPheSerGlyLeuTyrTrpThrLeuAlaLeuPheCysTrpIleSerAspArga 197
Db 17059 TGGGTTGCTGTCCTCCCT-----ACCTTTTGTGTTTTTT 17093
QY 197 laPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleL 217
Db 17094 --TTCTCTTTTCTGTACATAATTGACTACCT-----ATTA 17130
QY 217 euIleCysLeuAlaAlaTyr-----LeuGlyC 226
Db 17131 TACTATGCTCTTTCTTTTCTTCTACTTCTTTTAGTTTCTGCAAAACCCCTTTTCT 17190
QY 226 ysValCysPheAlaTyrPheAspAlaAlaSerGlu-----IleProGluGlnG 242
Db 17191 CTTTGTGTTGTCATATATCATGAAGGCAAAATCAATGTTCTCATCTTAGTACCCCTCAG 17250
QY 242 lyProValIleLysPheTrp 248
Db 17251 GGCCTGACACTCTGTTTTTC 17270

```

RESULT 5

```

US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereco
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature

```

LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Alignment Scores:

Pred. No.: 3.77 Length: 246240
Score: 114.00 Matches: 74
Percent Similarity: 33.44% Conservative: 29
Best Local Similarity: 24.03% Mismatches: 95
Query Match: 7.58% Indels: 110
DB: 2 Gaps: 14

US-10-017-410-4 (1-275) x US-08-724-394A-22 (1-246240)

Qy 13 GlySerSerGluValAspTrpCysGluAspAsnThrThrIleValProAlaIleAlaGlu 32
Db 16465 GGGAAATCGCTCACTGATTTTAAACAGAGACAGATGGTGTAGCTGTTTAAAAAGTAT 16524
Qy 33 PheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeu 52
Db 16525 TTTATAACCTTTAAAGAAATGCGGTTTGTCTCTCTTAAATCAATA---TTGTGCTTA 16581
Qy 53 PheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIle-TirThrLeuLeuVa 72
Db 16582 TGAGACCAATTC-----ATCTATTATATATTTGGGCTGTGTTAAT 16620
Qy 72 lValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLe 92
Db 16621 TCTCATTCGTATACAAATAGACCGTGGTGTATTTACTCATCTTTTAAACAGACAACT 16680
Qy 92 uAspGluLeuAla-----ValLeuTirPva 100
Db 16681 TCCTGTTGAGCTATTACAGATAGTCTCCTATGAACATTGTTGTACATGTTTATGAT 16740
Qy 100 lLeuMetCysAla-LeuAlaMetTirPheProArg-----A 112
Db 16741 GTATATGTCATTTCTAAAGTAAATGTTTGATGCTACTGGAATGTACATTTAGCTTTAGG 16800
Qy 112 rgTyrLeuProIlePheArgAsnAspArgGlyArgPheIysValValValSerValL 132
Db 16801 AGATATTCGCGAGCAGCTTT-----CCAAAGTTCACATTCGTAAAGTAAAGAA 16848
Qy 132 euSerAlaValThrThrCysLeuAla-----PheValIysProAlaIleAsnAsnI 149
Db 16849 GTTCAGTGTCTCCACATGTCGGCAACATTTGTTTCCATCTCTTCACTTTATATCCA 16908
Qy 149 leSer-----LeuM 152
Db 16909 TTAGTTAATTTTATAGTGGTATTTTCATTTGATGTTTGTATTTTCATTTCTGTAATGACTAA 16968
Qy 152 etThrLeuGlyVal-----P 157
Db 16969 TGAAGTTGAACATCTTTTATATGTTAAGGAGCCACTTACATTTCTCTTTTATGAAGTC 17028
Qy 157 rcCysThrAlaLeuLeuIleAlaGluLeuIysArgCysAspAsnMetArgValPheLysL 177
Db 17029 CCTGTTCAAGT-----CATTCGCGCCATTTTCTTACT 17058
Qy 177 euGlyLeuPheSerGlyLeuTirPirPirThrLeuAlaLeuPheCysTirPirIleSerArgA 197
Db 17059 TGGGTGTCTGTCGCCCC-----ACCTTTTGTGTTT----- 17093
Qy 197 laPheCysGluLeuLeuSerPheAsnPheProTyrLeuHisCysMetTirPheIleLeuL 217
Db 17094 --TTTCTCTTTTCTGTACATAAATTTGACTACCT-----ATTA 17130
Qy 217 euIleCysLeuAlaAlaTyr-----LeuGlyC 226
Db 17131 TACTATGCTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 17190
Qy 226 ysValCysPheAlaTyrPheAspAlaAlaSerGlu-----IleProGluGlnG 242
Db 17191 CTTTGTGTGTCATATCATGAGGCAAAATCAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 17250

Qy 242 lyProValIleLysPheTirP 248
Db 17251 GGCTGACACTCTGTTTTC 17270

RESULT 6

US-09-540-236-938
Sequence 938, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS

FILE REFERENCE: 2709.2005-001

CURRENT APPLICATION NUMBER: US/09/540,236

CURRENT FILING DATE: 2000-04-04

NUMBER OF SEQ ID NOS: 3840

SEQ ID NO 938

LENGTH: 1845

TYPE: DNA

ORGANISM: M. catarrhalis

US-09-540-236-938

Alignment Scores:
Pred. No.: 0.0395 Length: 1845
Score: 104.50 Matches: 69
Percent Similarity: 40.17% Conservative: 27
Best Local Similarity: 28.87% Mismatches: 79
Query Match: 6.95% Indels: 65
DB: 4 Gaps: 17

US-10-017-410-4 (1-275) x US-09-540-236-938 (1-1845)

Qy 34 TyrAsnThrIleSerAsnVal-----LeuPhePheIleLeu-----ProIleCys 49
Db 14 TTTCTATTGTTGAAACAGTTTAGAGCTCTTTCTGTATTTAGCTGCTCCACCATACCC 73
Qy 50 MetCysLeuPheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTirThr 69
Db 74 ATTCG-CTGGCTTCATTTATTTATGCTGCTGTTGGGA-----TATCGTGGTGGCTT 126
Qy 70 LeuLeu-----ValValValGlyIleGlySerValTyrPheHisPheThrLeuSerPhe 87
Db 127 TTGTTGTTTACAGTATTAAACAGCAATTGACGATCATTTGGGCATATGTGTATCGGTGG 186
Qy 88 LeuGlyGlnMetLeuAspGluLeuAlaVal-----LeuTirP 99
Db 187 ATTGGGCAATTTGTTGAGCTGCTGCTAGTGTATTTCTCCAGTGAATTTGGACACAAAA 246
Qy 100 -----ValLeuMetCysAlaLeuAlaMetTirPheProArgTyrLeuPro 115
Db 247 AAGTCAGATCTATTGATCATGCTGGCAATCAGTGCATTCATTCCTGTTGGTCTTTTG 303
Qy 116 LysIlePheArgAsnAspArgGlyArgPheIysValVal-----SerValLeuSer 133
Db 304 -----TTTGAACCAACCATTCACCATCAGGTGCTTACAA 336
Qy 134 AlaValThr-ThrCysLeu-----AlaPheValIysProAla-----Il 146
Db 337 GGAGTAGTACCAATCGTTTACGCTGGATTTTTCATCAGCAATGCTGGCCAGTCGATG 396
Qy 146 eAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLe 166
Db 397 CAATTTTATCAAGATGAATTTCTCTGCGGTG---TGTGAGCCAGGTGATGCAACCCGCT 453
Qy 166 uLysArgCysAspAsnMetArg-----ValPheIysLeuGlyLeuPheSerClyLe 183
Db 454 TTAGCGGTGCGT 513
Qy 183 uTirPirThrLeuAlaLeu-PheCysTirPirIleSer-----AspArgAlaPheCysG 200
Db 514 TTGGTGGGAGCAGGGGCTGTTTGTGCTGGATTTACCCACTACTGACAGTGCATTTGTA 573

```
QY 200 luleu---LeuSerSerPheAsnPhProTyTrLeuHisCysMetTrpHisIleLeuLeu 219
Db 574 TTTTGGATTAGCGGGTTTCGGCTTT-----ATTGGATT 609
QY 219 ysluAla-----AlaTyTrLeu-----GlyCysValCysPhe 229
Db 610 GTATTGCCAGCCTTCAGCTACAGTCGGATACAGCGGATGCTCGTCTTT 662

RESULT 7
US-09-596-002-41
; Sequence 41, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 269223
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 41
; PUBLICATION INFORMATION:
US-09-596-002-41

Alignment Scores:
Pred. No.: 87.6 Length: 269223
Score: 102.50 Matches: 61
Percent Similarity: 39.44% Conservatives: 23
Best Local Similarity: 28.64% Mismatches: 68
Query Match: 6.82% Indels: 61
DB: 4 Gaps: 15

US-10-017-410-4 (1-275) x US-09-596-002-41 (1-269223)
QY 56 TyrAlaThrCysLeuAsnSerAspIleTyTrLeuIleTrpThrLeu-----ValVal 73
Db 144414 TATGCTGTACTTTGGGA-----TATCGTGGTTGGCTTTTGTATTACAGTATTA 144464
QY 74 ValGlyIleGlySerValTyTrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuApp 93
Db 144465 ACAGCAATTCAGGCATCTATTGGCATATGTGTATGCGTGGATTCGGCAAAATTGTGAC 144524
QY 94 GluLeuAlaVal-----LeuTrp-----ValLeu 101
Db 144525 TGGCTTAGTGTGATTCTCCAGTGAATTTGGGCACAAAAAATCAGATCTATTGATC 144584
QY 102 MetCysAlaLeuAlaMetTrpPheProArgTyTrLeuProLysIlePheArgAsnApp 121
Db 144585 ATGCTGGCAATCAGTCGATTCATTTCGGCTGGTCTTTG-----144623
QY 122 ArgGlyArgPheLysValValVal-----SerValLeuSerAlaValThr-ThrCysLe 139
Db 144624 -----TTTGAGACACCACTTACCATCAGGTACTACAGGGGTATGCGCAATGCGT 144674
QY 139 u-----AlaPheValLysProAla-----IleAsnAsnIleSerLeuMe 152
Db 144675 TTACGTGGATTTTTCATCAGCAATGCTGGGCCAGTCGATGCAATTTTATCAAGATGA 144734
QY 152 tThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMe 172
Db 144735 TTCTCTGGCGGTG-----TGTACGCCAAGGTGATGCCAACTGCTTTAGCGGTGGTGATGTG 144791
QY 172 tArg-----ValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAlaLe 189
```

```
Db 144792 GTCATGACGCTGATTGGTGTGTGTTTCTTTATTCTCACAATTTTGTGGAGCAGCGGCT 144851
QY 189 u-PheCysTrpIleSer-----AspArgAlaPheCysGluLeu---LeuSerSerP 205
Db 144852 GTTTTGTAGTGGGATTCCACCACACTACTACAGAGTGCCATTTGTATTGTGGATTTTAGCGGTT 144911
QY 205 heAsnPhProTyTrLeuHisCysMetTrpHisIleLeuLeuCysLeuAla-----A 222
Db 144912 TCGGCTTT-----ATTGTATTGTATTGTCCACCGCTTCAG 144947
QY 222 latTyTrLeu-----GlyCysValCysPhe 229
Db 144948 CGTACCAGTCGGATACAGCGGATGCTCGTCTTT 144982

RESULT 8
US-09-252-991A-892
; Sequence 892, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 892
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-892

Alignment Scores:
Pred. No.: 0.0649 Length: 1257
Score: 100.50 Matches: 57
Percent Similarity: 38.43% Conservatives: 36
Best Local Similarity: 23.55% Mismatches: 59
Query Match: 6.69% Indels: 91
DB: 4 Gaps: 12

US-10-017-410-4 (1-275) x US-09-252-991A-892 (1-1257)
QY 46 ProProIleCysMetCysLeuPheAspGluTyTrAlaThr-----58
Db 483 CCGTCTACGTGCGGCTGTAC-ITCAGCACCTACGGCGCAGCGCGCCAGCCAGGAAACC 541
QY 59 -----CysLeuAsnSer 62
Db 542 CTGATGGGCTTCAACGGGGCTTCTACGGGGTGTCTTTCGGGTGTGTCTGTACAAACCTG 601
QY 63 AspIleTyTrLeu-----IleTrpThrLeuLeuValValGly 75
Db 602 TTCTCTTCTGCTCGCTACGCGAGGCCACCTATGCTGTGTCTGTACAACTCAGC 661
QY 76 IleGlySerValTyTrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp-----93
Db 662 CTCGGCCTG-----TTTTCGGCAGCTTCAGCGGCTGTGTTCAGGCTGTG 709
QY 94 -----GluLeuAlaValLeuTrpValLeuMet-----CysAla 104
Db 710 CCGGACAGTGGCGCTGAGTCGGCGGCATCTACCTGCTGTATGCTAGTGCCTGCTG 769
QY 105 LeuAlaMetTrpPheProArgArgTyTrLeu-----ProLysIlePhe 118
Db 770 GTGTGCTCAGTTCAGCGCGGCTACCTCTACACCGCGCGGACTTCCCGCGCTCGAC 829
QY 119 ArgAsnAspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCys 138
```

830	CGCTTCCTCCGGCGCTGCTGCTGGCGCTGCTGGTCTCTGTGGCCAGC-----	877
139	LeuAlaPheValIysProAlaIle-----AsnAsnIleSerLeuMetThr	153
878	-----GAGCCGCTGGTGGGGCTGCGCGCTGGAAAGTCTCGCCAGCCTGACG	925
154	LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuIysArgCysAspAsnMetArg	173
926	GTGATGCTGCTCTCCCTCAGCCTGTCTGTGGCC-----GGCGTCCAC	967
174	ValPheLysLeuGluGluPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIle	193
968	GTCTGGCGCCAGGGCTTGGCTACGGCCTCTACTACATCTGGCC-----TGGGGC	1018
194	SerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMet	213
1019	GCG-----CTGCTGCTCTCGTTC-----	1036
214	TrpHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAsp	233
1037	-----CTGTGCACCCAGCGCGCTCGCTGGGCTGCGAATGTTTCGGCGTCTTCGC	1087
234	AlaAla	235
1088	AGCAGC	1093

```

RESULT 9
US-09-252-991A-944
; Sequence 944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 944
; LENGTH: 2859
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-944

```

Alignment Scores:		
Pred. No.:	0.212	2859
Score:	100.50	57
Percent Similarity:	38.43%	36
Best Local Similarity:	23.55%	59
Query Match:	6.63%	91
DB:	4	12
US-10-017-410-4 (1-275)	x	US-09-252-991A-944 (1-2859)

Qy	46	ProProlleCysMetCysLeuPheAspGluTyrAlaThr	-----	58
Db	557	CCGTCTACGTGCGGTGTAC..TTCAGCACCTACGCGCCAGCGCGGCCAGCAGAAACC	615	
Qy	59	-----CysLeuAsnSer	62	
Db	616	CTGATGGGCTTCAACGGGGCCTTCTACGGGGTGTGTTTCGGTGTCTGTACAACTCG	675	
Qy	63	AspIleTyrLeu-----lleTprThrLeuLeuValValValGly	75	
Db	676	TTCCTCTTCGCTCGCTACGCGAGGCCACTATGCTGTGTACCTGCTGTACAACTCAGC	735	
Qy	76	lleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp	93	
Db	736	CTCGGCGCTG-----TTTTCGGACAGCTTCGACGGCGCTGCTGTCAAGCTGCTGT	783	

Qy	94	-----GluLeuAlaValLeuTrpValLeuMet-----CysAla	104
Db	784	CCCGCACACGTGGCGCTGCAGTCGGCGCATCTTACCTGCTGATGATTACTGAGTGCGCTG	843
Qy	105	LeuAlaMetTrpPheProArgArgTyrLeu-----ProLysilePhe	118
Db	844	GTTGTCGATCCAGTTTCAGCGCGGCTACCTCTACACC CGCGCGCACTTCCCGCGCTCGAC	903
Qy	119	ArgAsnAspArgGlyArgPheLysValValSerValLeuSerAlaValThrCys	138
Db	904	CGCTTCTCCGCGGCGCTGCTGCCTGCCCTGCCTGGTGCCTGTGGCGCAGC-----	951
Qy	139	LeuAlaPheValLysProAlaIle-----AsnAsnIleSerLeuMetThr	153
Db	952	-----GAGCGCTGTGGGGCTGCGCGCTGGAACGCTCTCGCCACGCTGACG	999
Qy	154	LeuGlyValProCysThrAlaLeuLeuLeuAlaGluLeuLysArgCysAspAsnMetArg	173
Db	1000	GTTGATCTGCTGCCCTCCTCAGCCTGCTGTGGCC-----GGCGTCCAC	1041
Qy	174	ValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIle	193
Db	1042	GTCGTGGCGCAGCGCTGCGCTACGCGCCCTCTACTACATCTCGCC-----TGGGCG	1092
Qy	194	SerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMet	213
Db	1093	GCG-----CTGCTGCTGCTGTTTC-----	1110
Qy	214	TrpHisIleLeulleCysleuAlaIatYrLeuGlyCysValCysPheAlaTyrPheasp	233
Db	1111	-----CTGTGCACCAACCGCGCTCGCTGGGTGCGAACTGTTCGGCTGTTCGGC	1161
Qy	234	AlaAla	235
Db	1162	AGCAGC	1167
 RESULT 10			
US-09-252-991A-1139/c			
; Sequence 1139, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS			
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 1139			
; LENGTH: 1575			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-1139			

Alignment Scores:		
Pred. No.:	0.102	Length: 1575
Score:	100.00	Matches: 49
Percent Similarity:	43.16%	Conservative: 33
Best Local Similarity:	25.79%	Mismatches: 48
Query Match:	6.65%	Indels: 10
DB:	4	Gaps: 10

US-10-017-410-4 (1-275) x US-09-252-991A-1139 (1-1575)

Qy 68 TrpThrLeuLeuValValGlyIleGlySerValTyrPheHisPheThrLeuSerPhe 87
||| ||| ||| :|: :|: ||| :|: ||| :|:
Db 1554 TGGACCTGCTGTACAACCTCAGCTCGGCTG-----TTTTCGCCAGCTTC 1507

[illegible]

QY 88 LeuGlyGlnMetLeuasp-----GluLeuAlaValLeuTyr 99
Db 1506 GAGGGCTGCTGTTAAAGCTGCTGCCGGACAGTGGCGCTGGAGTCGGCGCATCTAC 1447
QY 100 ValLeuMet-----CysAlaLeuAlaMetTyrPheProArgArgTyrLeu----- 114
Db 1446 CTGCTGATGACCTGAGCTGCTGGTGTGATCCAGTTACAGCCGGCTACCTCTACACC 1387
QY 115 -----ProLysPhePheArgAsnAspArgGlyArgPheLysValValSer 130
Db 1386 CGCCGGAGCTTCGCCGCTCGACCGCTTCTCCGGCGCTGCTGGCGCTCGGTGTC 1327
QY 131 ValLeuSerAlaValThrThrCysLeuAlaPheValLysProAlaIle----- 146
Db 1326 CTGTTGGCCAGC-----GAGCCGCTGGTGGCGCTGGCGCC 1291
QY 147 ---AsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuAlaGlu 165
Db 1290 TGAACGCTCCCGCAGCTGACGGTGTCTGCTCCCTCAGCCTGCTGCTGGCC--- 1234
QY 166 LeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTyrTyr 185
Db 1233 -----GCGCTCCACGCTGGCGCCAGCGCTGCGCTACGSCCTCTACTAC 1189
QY 186 ThrLeuAlaLeuPheCysTyrPheSerArgAlaPheCysGluLeuLeuSerPhe 205
Db 1188 ATCTGCGC-----TGGCGCGC-----CTGCTGCTGCTGCTG 1156
QY 206 AsnPheProTyrLeuHisCysMetTyrHisIleLeuLeuLeuAlaAlaTyrLeuGly 225
Db 1155 -----CTGTCACACCCCGCTGCTGGCG 1129
QY 226 CysValCysPheAlaTyrPheAspAlaAla 235
Db 1128 TCGGAAGCTGTGGCGCTGTTGGCGCAGCAGC 1099

RESULT 11

US-09-252-991A-15016
; Sequence 15016, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15016
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15016

Alignment Scores:

Pred. No.: 0.163 Length: 1260
Score: 97.00 Matches: 56
Percent Similarity: 35.90% Conservatives: 28
Best Local Similarity: 23.93% Mismatches: 87
Query Match: 6.45% Indels: 63
DB: 4 Gaps: 10

US-10-017-410-4 (1-275) x US-09-252-991A-15016 (1-1260)

QY 19 TrpCysGluAsnTyrThrIleValProAla---IleAlaGluPheTyrAsnThrIle 37
Db 374 TGGACCCGGAGTGTACGAAGCGCCCGCGCGCTGCTCGCGCTTCTCGCCACCGCC 433
QY 38 SerAsnValLeuPhePheIleLeuProLleCysMetCysLeuPheAspGluTyrAla 57

Db 434 AGCAAGTCTCGCGTGTTC-----GCGTCTGCTGCGGCTGTTCAGATCGCCCG 484
QY 58 ThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuLeuValValGlyIleGly 77
Db 485 GCGGCCCTCGACCAACCACTGTCTGAACATC-----TCCCTGAGCGTCTATCGCGTCC 538
QY 78 SerValTyrPhe-----HisPheThrLeuSerPheLeuGlyGlnMetLeu--- 92
Db 539 TCGATCTCTTCGGCAACCTGTGGCACTGACCCAGAGCAACATCAAGCCCTGCTCGGC 598
QY 93 ---AspGluLeuAlaValLeuTyrValLeuMetCysAlaLeu----- 105
Db 599 TACTCGTCCATCGCCACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
QY 106 AlaMetTyrPheProArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPhe 125
Db 659 GCGCTGGAAGCGGTGCGGCTCTACTTGGCCACCTACGTACTGACCTCCCTCGCGCGTTC 718
QY 126 LysValValVal----- 129
Db 719 GCGGTGATCACCTGATGTCCACCCGCTACAGCGCGCGGCGGCGGCGGCGGCGGCGG 778
QY 130 -----SerValLeuSerAlaValThrThrCysLeuAlaPhe 141
Db 779 TACCGCGGCTGCTTCTGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832
QY 142 ValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeu 161
Db 833 -----ATGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 868
QY 162 LeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSer 181
Db 869 TTCATCGGCAAG-----TTCACGTGATCGCGTCTGCGGCTGCGGCTGCGGCTGCG 910
QY 182 GlyLeuTyrThrThrLeu-AlaLeuPheCysTyrPheSerAspArgAlaPheCysGluLe 201
Db 911 CACCTGTGTGCTGATCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 970
QY 201 uleuSerSerPheAsnPheProTyrLeuHisCysMetTyr 214
Db 971 CTGCGGCTCATGTGTCA-----CCCTGTTCTCTGG 998

RESULT 12

US-09-252-991A-15148
; Sequence 15148, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15148
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15148

Alignment Scores:

Pred. No.: 0.231 Length: 1605
Score: 97.00 Matches: 56
Percent Similarity: 35.90% Conservatives: 28
Best Local Similarity: 23.93% Mismatches: 87
Query Match: 6.45% Indels: 63
DB: 4 Gaps: 10

US-10-017-410-4 (1-275) x US-09-252-991A-15148 (1-1605)

```
QY 19 TrpCysGluAspAsnTyrThrIleValProAla-----IleAlaGluPheTyrAsnThrIle 37
Db 820 TGGACCCCGGATGTCTACGAAGCGCGCGCGCGGTCTCGCGGTTCCTCGCCACCGCC 879
QY 38 SerAsnValLeuPhePheIleLeuProPheIleCysMetCysLeuPheAspGluTyrAla 57
Db 880 AGCAAGTCCGGTGTTC-----GCCGTGCTCTCGCGGTGTTCAGATCGCCCGC 930
QY 58 ThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuLeuValValGlyIleGly 77
Db 931 GCGGCCCTGGACACCACTGCTGAATC-----TCCCTGAGCGTCTCGCGGTCCG 384
QY 78 SerValTyrPhe-----HisPheThrLeuSerPheLeuGlyGlnMetLeu--- 92
Db 985 TCGATCCTCTCGGCAACCTGCTGGCACTGACCCAGACCAATCAAGCGCTGCTCGC 1044
QY 93 ---AspGluLeuAlaValLeuTyrValLeuMetCysAlaLeu----- 105
Db 1045 TACTCGTCCATCGCCCACTGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104
QY 106 AlaMetTyrPheProArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPhe 125
Db 1105 GCGGTGGAAGCGGTGCGGTCTACCTGGCCACCTACGTACTGACCTCCCTCGCGCG 1164
QY 126 LysValValVal----- 129
Db 1165 GCGGTGATCACCCTGATGTCACCCCGTACAGCGCGCGCGCGCGATGCGGTTCGAG 1224
QY 130 -----SerValLeuSerAlaValThrCysLeuAlaPhe 141
Db 1225 TACCGGGCGCTGTTCTGGCGCGCGCGGTACTGACCGCGTGATGACCGGTGATG 1278
QY 142 ValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeu 161
Db 1279 -----ATGCTGCTGCTGCGAGGCATCCCGCTGACCGCGCGCGCGCGCG 1314
QY 162 LeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSer 181
Db 1315 TTATCGGCAAG-----TTCTACGTGATCGCGGTGCGGTGCGGTGCGGTGCG 1356
QY 182 GlyLeuTyrThrThrLeu-AlaLeuPheCysTyrIleSerAspArgAlaPheCysGluLe 201
Db 1357 CACTGTGGTGGTGTGATCGCGCGCGCTGCTGCTGCGGAGCCATCGCGCTGTACTAC 1416
QY 201 uLeuSerSerPheAsnPheProTyrLeuHisCysMetTyr 214
Db 1417 CTGCGGTGATGGTCA-----CCCTGTTCCTGG 1444
```

RESULT 13

US-09-252-991A-14616/c
; Sequence 14616, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: 1999-02-18
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14616
; LENGTH: 1950
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14616
Alignment Scores:

Pred. No.: 0.306 Length: 1950
Score: 97.00 Matches: 56
Percent Similarity: 35.90% Conservative: 28
Best Local Similarity: 23.93% Mismatches: 87
Query Match: 6.45% Indels: 63
DB: 4 Gaps: 10

US-10-017-410-4 (1-275) x US-09-252-991A-14616 (1-1950)

```
QY 19 TrpCysGluAspAsnTyrThrIleValProAla-----IleAlaGluPheTyrAsnThrIle 37
Db 1056 TGGACCCCGGATGTCTACGAAGCGCGCGCGCGGTCTCGCGGTTCCTCGCCACCGCC 997
QY 38 SerAsnValLeuPhePheIleLeuProPheIleCysMetCysLeuPheAspGluTyrAla 57
Db 996 AGCAAGTCCGGTGTTC-----GCCGTGCTCTCGCGGTGTTCAGATCGCCCGC 946
QY 58 ThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuLeuValValGlyIleGly 77
Db 945 GCGGCCCTGGACACCACTGCTGAATC-----TCCCTGAGCGTCTCGCGGTCCG 892
QY 78 SerValTyrPhe-----HisPheThrLeuSerPheLeuGlyGlnMetLeu--- 92
Db 891 TCGATCCTCTCGGCAACCTGCTGGCACTGACCCAGACCAATCAAGCGCTGCTCGC 832
QY 93 ---AspGluLeuAlaValLeuTyrValLeuMetCysAlaLeu----- 105
Db 831 TACTCGTCCATCGCCCACTGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772
QY 106 AlaMetTyrPheProArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPhe 125
Db 771 GCGGTGGAAGCGGTGCGGTCTACCTGGCCACCTACGTACTGACCTCCCTCGCGCG 712
QY 126 LysValValVal----- 129
Db 711 GCGGTGATCACCCTGATGTCACCCCGTACAGCGCGCGCGCGATGCGGTTCGAG 652
QY 130 -----SerValLeuSerAlaValThrCysLeuAlaPhe 141
Db 651 TACCGGGCGCTGTTCTGGCGCGCGGTACTGACCGCGTGATGACCGGTGATG- 598
QY 142 ValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeu 161
Db 597 -----ATGCTGCTGCTGCGAGGCATCCCGCTGACCGCGCGCGCGCG 562
QY 162 LeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSer 181
Db 561 TTATCGGCAAG-----TTCTACGTGATCGCGGTGCGGTGCGGTGCGGTGCG 520
QY 182 GlyLeuTyrThrThrLeu-AlaLeuPheCysTyrIleSerAspArgAlaPheCysGluLe 201
Db 519 CACTGTGGTGGTGTGATCGCGCGCGCTGCTGCTGCGGAGCCATCGCGCTGTACTAC 460
QY 201 uLeuSerSerPheAsnPheProTyrLeuHisCysMetTyr 214
Db 459 CTGCGGTGATGGTCA-----CCCTGTTCCTGG 432
```

RESULT 14

US-09-252-991A-2596
; Sequence 2596, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: 1999-02-18
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```
; SEQ ID NO 2596
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2596

Alignment Scores:
Pred. No.: 0.336 Length: 2082
Score: 97.00 Matches: 74
Percent Similarity: 35.43% Conservative: 50
Best Local Similarity: 21.14% Mismatches: 102
Query Match: 6.45% Indels: 125
DB: 4 Gaps: 18

US-10-017-410-4 (1-275) x US-09-252-991A-2596 (1-2082)

QY 2 GlyAlaProHisTrpAspGlnLeuGlnAlaGlySerSerGluVal----- 17
Db 979 GGCTCGGTACTGGATGCTCCACCTCGAGGCCCGCGCGCGGATCTAAAGGCGCTGCCG 1038
QY 18 ----AspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThr 36
Db 1039 CGCGATCAGTGTCCGAGCGCTGG-----GGTCCGCTCCGCGAGCGCTGG----- 1083
QY 37 IleSerAsnValLeuPhePheIleLeuProProIleCysMet-----CysLeuPheAsp 54
Db 1084 -----TTCTGTGCTGATTCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTCC 1128
QY 55 GluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuLeuValValVal 74
Db 1129 GGGCGACCGCGCTGTTCCGGAACCATCGGCTGGCGCTGACCGCCATCGTCACTCC 1188
QY 75 GlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGlu 94
Db 1189 GGTTCGGCGATCATC-----CTGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1233
QY 95 LeuAlaValLeuTrpVal-----LeuMetCys----- 103
Db 1234 -----GCCTTCGTGGTGGCGTGGGCGTCTCGCGGGGTTCTTCGCGCTCGGCATC 1287
QY 104 -----AlaLeuAlaMetTrpPheProArg----- 111
Db 1288 GGGGTGATCTTCGTGCTGATCGCGCCCTGGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTG 1347
QY 112 -----ArgTyrLeuProIlePheArgAsn----- 120
Db 1348 CGGAAACCTCGCATCTGCTCGCGCCCTCGCGAAGGTGCGGACATCGGATTCG 1407
QY 120 ----- 120
Db 1408 GTGGCATCGCTGCACCTGGTGGAGTCATCATCGGGGTGATTTCGCTGACCGGCATC 1467
QY 121 -----AspArgGlyArgPheLysValValValValValValValValValValVal 131
Db 1468 GCCACCACTTCGCCAGCTACGTGCTGACCTGGGACGGAGAACCTGCTGCTGCTGCTGCTG 1527
QY 132 LeuSerAlaValThrCysLeuAlaPheValLys-----ProAlaIleAsnAsnIle 149
Db 1528 GTCTGACCATGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1587
QY 150 -----SerLeuMetThrLeuGlyValProCysThrAla 160
Db 1588 ATCATCACCAGCTCCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db 1639 CTGATCGTCTCG-----CACATGTTCTGCTTCTACTTCTGCGCATCTC 1680
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db 1681 GCCACCTCAACCTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpPheIleLeuLeuLeuLeu 220
```

```
1741 AGGGATTGAAGATCAGCTTC-----TGGCGGTGCGCATCGCCCTG 1782
221 AlaAlaTyrLeuGlyCysValCysPhe-----AlaTyrPheAspAlaAlaSerGluIlePro 239
1783 GCGGTTTCGTG-----GTGCGGTTTCATGCGCGGTGATTTCCCGCGCTGATGCTGCAG 1836
240 GluGlnGly-----ProValIleLysPheTrpProAsnGluLys 252
1837 GAAGCGGTTGGGGGCGGAGCTCTATGTCACCTGCAAGCGGTGCTGGCCATCGGCGCTG 1896
252 sTrpAlaPheIleGlyValProTyrVal 261
1897 -TGGGCGATGGCTTCATCGCTACCTG 1923

RESULT 15
US-09-252-991A-2514
; Sequence 2514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2514
; LENGTH: 2388
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2514

Alignment Scores:
Pred. No.: 0.41 Length: 2388
Score: 97.00 Matches: 74
Percent Similarity: 35.43% Conservative: 50
Best Local Similarity: 21.14% Mismatches: 102
Query Match: 6.45% Indels: 125
DB: 4 Gaps: 18

US-10-017-410-4 (1-275) x US-09-252-991A-2514 (1-2388)

QY 2 GlyAlaProHisTrpAspGlnLeuGlnAlaGlySerSerGluVal----- 17
Db 890 GGCTCGGTGTAATGATGCTCCACCTCGAGGCCCGCGCGCGATCTAAAGGCGCTGCCG 949
QY 18 ----AspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThr 36
Db 950 CGCATCAGTGTCCGAGCGCTGG-----GGTCCGCTCCGCGAGCGCTGG----- 994
QY 37 IleSerAsnValLeuPhePheIleLeuProProIleCysMet-----CysLeuPheAsp 54
Db 995 -----TTCTGTGATTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039
QY 55 GluTyrAlaThrCysLeuAsnSerAspIleTyrIleTyrThrLeuLeuValValVal 74
Db 1040 GGGCGACCGCGCTGTTCTCCGAACCATCGGCTGGCGCTGACCGCATCGTCACTC 1099
QY 75 GlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGlu 94
Db 1100 GGTTCGGCGATCATC-----CTGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1144
QY 95 LeuAlaValLeuTrpVal-----LeuMetCys----- 103
Db 1145 -----GCCTTCGTGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198
QY 104 -----AlaLeuAlaMetTrpPheProArg----- 111
```


Db 1199 GGGGTGATCTTCGTGATCGCGGCCCTGGCCCTGGCTGCTGTTGCTGCTGCGGC 1258
Qy 112 -----ArgTyrLeuProLysIlePheArgAsn----- 120
Db 1259 CGGGAACCTTCGGCATCTCGTCCGCGCCCTCGCCGAAGTGGCGACATCGATTCCG 1318
Qy 120 ----- 120
Db 1319 GTGGGATCGCCTGCACCCCTGTGGAGTCAATCGGGGTGATTTCGTGACCGGCATC 1378
Qy 121 -----AspArgGlyArgPheLysValValSerVal 131
Db 1379 GCCACCACTTCGCCAGCTACGTGCTCGACCTGGGACGGGAGAACCTGTCTGTGCTG 1438
Qy 132 LeuSerAlaValThrThrCysLeuAlaPheValLys-----ProAlaIleAsnAsnIle 149
Db 1439 GTCTGACCATGTGCACCTGCTGCTGGGAATGGGCATCCCGACCATCCCAACTAC 1498
Qy 150 -----SerLeuMetThrThrLeuGlyValProCysThrAla 160
Db 1499 ATCATCACAGCTCCATCGCGCGCGCTGCTGGAACCTGGGGTGGCG----- 1549
Qy 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db 1550 CTGATCGTCTCG-----CACATGTTCTCTTCTACTTGGGCATCCTC 1591
Qy 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db 1592 GCGGACCTCACCCCTCGCGTGGCGCTGGCTTTTCGCGCGCGCGCGATCGCCCAAGAA 1651
Qy 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db 1652 AGGGATTGAAGATCAGCTTC-----TGGGCGGTGGCATCGCCCTG 1693
Qy 221 AlaAlaTyrLeuGlyCysValCysPhe---AlaTyrPheAspAlaAlaSerGluIlePro 239
Db 1694 GCGGGTTTCGTG-----GTGCGGTTTCATGCGGGGTGTATTCCCGCGGCTGATGCTGCAG 1747
Qy 240 GluGlnGly-----ProValIleLysPheTrpProAsnGluLys 252
Db 1748 GAAGCGGTTGGGGGGGCGACGCTATGTACCTGCAAGCGGTGCTGGCCATCGGCTG 1807
Qy 252 sTrpAlaPheIleGlyValProTyrVal 261
Db 1808 -TGGGCATGGCTTCCATCGGCTACCIG 1834

Search completed: September 18, 2004, 07:01:23
Job time : 275 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 05:15:15 ; Search time 444 Seconds
(without alignments)
3126.234 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 1503
Sequence: 1 MGAHPHWDQLQAGSEVDWC.....IGVPYVLLCANKKSSVKIT 275

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3327077 segs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPto.spool.p/US10017410/runat_15092004_164710_4747/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10017410 @CGN 1 1 723 @runat_15092004_164710_4747
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1503	100.0	828	14	US-10-017-410-3
2	1473	98.0	4202	17	US-10-188-832-21
3	1473	98.0	4212	16	US-10-295-027-151
4	1404	93.4	4175	14	US-10-017-410-1
5	1390.5	92.5	822	15	US-10-182-447-6
6	645	42.9	744	10	US-09-945-527-48
7	538.5	35.8	792	15	US-10-182-447-4
8	216.5	14.4	801	15	US-10-182-447-5
9	216.5	14.4	1063	13	US-10-302-172-875
10	206.5	13.7	1194	13	US-10-425-114-32295
11	202.5	13.5	1547	17	US-10-767-701-12261
12	198.5	13.2	1436	17	US-10-437-963-42300
13	186	12.4	35425	15	US-10-017-161-2429
14	186	12.4	35425	16	US-10-292-798-2069
15	178	11.8	1357	13	US-10-424-599-34549
16	159	10.6	636	16	US-10-264-237-615
17	123.5	8.2	2186	13	US-10-425-114-33305
18	119	7.9	248436	13	US-10-087-192-2014
19	118	7.9	568	17	US-10-767-701-30658
20	115	7.7	284	15	US-10-002-631C-17
21	114	7.6	235033	15	US-10-301-844-1
22	111.5	7.4	1498	15	US-10-017-161-1999
23	111.5	7.4	1498	16	US-10-292-798-1645
24	111	7.4	5842	11	US-09-984-429-316
25	110	7.3	1501	15	US-10-017-161-2097
26	110	7.3	1501	16	US-10-292-798-1743
27	108.5	7.2	1984	15	US-10-017-161-2133
28	108.5	7.2	2614	16	US-10-292-798-1779
29	107	7.1	237326	15	US-10-301-844-2
30	102.5	6.8	269223	13	US-10-672-781-41
31	102	6.8	81940	9	US-09-759-508B-1
32	102	6.8	81940	10	US-09-960-706-1092
33	101.5	6.8	1282	15	US-10-017-161-2059
34	101.5	6.8	1282	16	US-10-292-798-1745
35	101.5	6.7	1186	15	US-10-017-161-2113
36	100.5	6.7	1186	16	US-10-292-798-1759
37	100.5	6.7	180557	14	US-10-003-806-6
38	100.5	6.7	180557	15	US-10-003-806-9
39	100.5	6.7	180557	16	US-10-292-798-1369
40	100.5	6.7	744802	13	US-10-282-122A-17482
41	100	6.7	2061	13	US-10-198-846-13523
42	100	6.7	2485	15	US-10-424-599-127764
43	99.5	6.6	757	13	US-10-198-846-14046
44	99	6.6	1952	15	US-09-770-445-596
45	98.5	6.6	753	9	US-09-770-445-596

Alignment Scores:

ALIGNMENTS

RESULT 1
US-10-017-410-3
; Sequence 3, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(825)
US-10-017-410-3

Pred. No.:	2.4e-162	Length:	828
Score:	1503.00%	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-017-410-4 (1-275) x US-10-017-410-3 (1-828)

QY	1	MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys	20
DB	1	ATGGGGCCCGCACCTGGTGGGACCAGCTGCAGGCTGTAGCTCGAGGTGGA	60
QY	21	GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal	40
DB	61	GAGGACAACTACACCATCGTGGCTCTATCCGCGAGTTCTACAAACAGCATCAGCAATGTC	120
QY	41	LeuPhePheIleLeuProProIleCysMetCysLeuLeuPheAspGluTyrAlaThrCysLeu	60
DB	121	TTATTTTTCATTTTACCGCCCATCTGCATGTGCTGTTGATGAGTAGATGCAACATGCTTG	180
QY	61	AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr	80
DB	181	AACAGTGAACATCTACTTAATCTGGACTCTTTTGGTGTAGTGGGAATGGATCCGCTAC	240
QY	81	PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal	100
DB	241	TTCCATTTTACCCTTAGTTTCTTGGTCAGATGCTTGATGAACCTTGCAGTCCCTTTGGGTT	300
QY	101	LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProIlySillePheArgAsn	120
DB	301	CTGATGTGTGCTTTGGCCATGTGGTTCGCCAAGAGGTATCTACCAAGATCTTCGGAT	360
QY	121	AspArgGlyArgPheLeuValValValSerValLeuSerAlaValThrThrCysLeuAla	140
DB	361	GACAGGGTAGGTTCAGGGTGGTGGTCAAGTGTCTGTCTGCGGTTCAGACGTGCCTGGCA	420
QY	141	PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla	160
DB	421	TTTGTCAAGCTGCCATCAACAAATCTCTCTGATGACCCCTGGAGTTCCTTGCACTGCA	480
QY	161	LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe	180
DB	481	CTGTCTCATCGCAGAGCTAAAGAGGTGTGCACAAATGCGTGTGTTTAAGCTGGGCTCTTC	540
QY	181	SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu	200
DB	541	TCGGGCTCTGGTGGACCTGGGCCCTGTTCTGTGGATCAGTGACCGAGCTTCTGCGAG	600
QY	201	LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu	220
DB	601	CTGTGTTCATCCTTCAACTTCCCTTACCTGCATCTGCATGTGGCACAATCCTCATCTGCCTT	660
QY	221	AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaIleSerGluIleProGlu	240
DB	661	GCTGCCCTACCTGGGCTGTATGCTTTGGCTACTTTTGATGCTGCTCTCAGAGATTCCTGAG	720
QY	241	GinglyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyr	260
DB	721	CAAGGCCCTGTTCATCAAGTTCTGGCCCAATGAGAAATGGGCCCTTCATGGTGTGCCCTAT	780
QY	261	ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr	275
DB	781	GTGTCCCTCTGTGTGCCAACAAAGAATATCATGATCAAGTCAAGTCAAG	825

RESULT 2

US-10-188-832-21
; Sequence 21, Application US/1018832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.

```

; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 4202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-21

```

Alignment Scores:

Pred. No.:	7.39e-158	Length:	4202
Score:	1473.00	Matches:	270
Percent Similarity:	98.55%	Conservative:	1
Best Local Similarity:	98.18%	Mismatches:	4
Query Match:	98.00%	Indels:	0
DB:	17	Gaps:	0

US-10-017-410-4 (1-275) x US-10-188-832-21 (1-4202)

QY	1	MetGlyAlaProHisTrrTpAspGlnLeuGlnAlaGlySerSerGluValAspTrrCys	20
Db	77	ATGGCGCGCCCGCACTGGTGGGACAGCTGCAGGCTGGTAGCTCGAGGTGGACTGGTGC	136
QY	21	GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal	40
Db	137	GAGGACAACACTACACATCGTGGCTGCTATCGCGAGTTCTACAAACACGATCAGCAATGTC	196
QY	41	LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu	60
Db	197	TTATTTTTCATTTTACCGCCCATCTGCATGGCTGTGTTTCGTCAATGATCAACATCGCTTC	256
QY	61	AsnSerAspIleTyrLeuIleTrrThrLeuLeuValValGlyIleGlySerValTyr	80
Db	257	AACAGTGCATCTACTTAACTCGGACTCTTTTGGTTAGTGGGAATTGGATCCGCTCTAC	316
QY	81	PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrrVal	100
Db	317	TTCCATGCAACCCCTTAGTTCTTTGGGTTCAGATGCTTGATGAACCTTCAGTCCCTTTGGGTT	376
QY	101	LeuMetCysAlaLeuAlaMetTrrPheProAtgArgTyrLeuProLysIlePheArgAsn	120
Db	377	CTGATGTGTGCTTTGGCCATGTGGTTCCTCCAGAAGGTATCTCAAGAAGATCTTTGGGAAT	436
QY	121	AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla	140
Db	437	GACCGGGGTAGTTTCAAGGTGGTGGTCAGTGTCTCTGCTCGGGTTACGACGTGCCCTGGCA	496
QY	141	PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla	160
Db	497	TTTGTCAAGCGCTGCCATCAACAACATCTCTCTGATGACCCCTGGGAGTTCCTTGCACCTGCA	556
QY	161	LeuLeuIleAlaGluLeuLysArgCysAspAsnMetAtgValPheLysLeuGlyLeuPhe	180
Db	557	CTGCTCATCGCAGAGCTAAGAGGTGTGACACATCGCGTGTGTTTAAAGCTGGGCCTCTTC	616
QY	181	SerGlyLeuTrrTrrThrLeuAlaLeuPheCysTrrPileSerAspArgAlaPheCysGlu	200
Db	617	TCGGGCTCTGGTGGACCCCTGGGCCCTGTTCTCTGGATCAGTGAACCGAGCTCTTCGAG	676

QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuLeuLeuCysLeu 220
Db 677 CTGCTGTCATCCCTCAACATTCCTCCCTACCTGCACTGATGGGACATCCCTCATCTGCTT 736
QY 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 737 GCTGCTACTCGGCTGTGTATGCTTTGCTACTTTGATGCTGCTCCAGAGATTCCTGAG 796
QY 241 GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyr 260
Db 797 CAAGGCCCTGTGTCATCAAGTTCTGGCCCAATGAGAAATGGCCCTTCATTTGGTGTCCCTAT 856
QY 261 ValSerLeuLeuCysAlaAsnLysSerSerValIleThr 275
Db 857 GTGTCCCTCTGTGTGTCACCAAGAAATCATCATGATCAAGATCAG 901

RESULT 3

US-10-295-027-151
; Sequence 151, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 4212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-151

Alignment Scores:
Pred. No.: 7 42e-158 Length: 4212
Score: 1473.00 Matches: 270
Percent Similarity: 98.55% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 4
Query Match: 98.00% Indels: 0
Gaps: 16

US-10-017-410-4 (1-275) x US-10-295-027-151 (1-4212)
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db 87 ATGGGCGCCCGCACCTGGTGGGACAGCTGCAGGCTGTAGCTCGAGGTGGACTGGTGC 146
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 147 GAGGACAACTACACCATCGTGCCTGCTATCGCCGAGTCTTACAAACAGATCAGCAATGTC 206
QY 41 LeuPhePheIleLeuProProfileCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db 207 TTATTTTTCATTTTACCGCCCATCTCATGTGCTTGTTCGTAGTATGCAATGCTTTC 266
QY 61 AsnSerAspIleTyrLeuIleTyrThrLeuValValValGlyIleGlySerValTyr 80
Db 267 AACAGTGGCACTACTTAACTCTGACCTCTTTGGTGTAGTGGGAATGGATCCGCTAC 326
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db 327 TTCCATGCAACCTTAGTTTCTTGGGTACAGATGCTTGATGAACCTTGAGTCTTGGGTT 386
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProTyrIlePheArgAsn 120
Db 387 CTGATGTGTGCTTTGGCCATGTGTTCCCGAAGGATATACCAAGATCTTTTCGGAAT 446
QY 121 AspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 447 GACCGGGTAGGTTCAAGGTGGTGTGCTGTCTGTCTGGGTACGAGCTGCTGGCA 506
QY 141 PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 507 TTGTCAAGCTGCATCAACAAACATCTCTGATGACCTGGGAGTTCCTTGCACCTGCA 566
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db 567 CTGCTCATCGCAGAGCTAAAGAGGTGTGACAAACATGGTGTGTAAAGCTGGGCTCTTC 626
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db 627 TCGGGCTCTGTGTGGACCCCTGGCCCTGTTCTGCTGGATCATGACCGAGCTTTCTGGAG 686
QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db 687 CTGCTGTCATCTTCAACTTCCCTACTGTCATGTCATGTGGCACATCCTCATCTGCCTT 746
QY 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 747 GCTGCTACTCGGCTGTGTATGCTTTGCTTACTTGTATGCTGCTCAGAGATTCCTGAG 806
QY 241 GlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyr 260
Db 807 CAAGCCCTGTGTCATCAAGTTCTGGCCCAATGAGAAATGGGCTTCATTTGGTGTCCCTAT 866
QY 261 ValSerLeuLeuCysAlaAsnLysSerSerValIleThr 275
Db 867 GTGTCCCTCTGTGTGTCACCAAGAAATCATCATGATCAAGATCAG 911
RESULT 4
US-10-017-410-1
; Sequence 1, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

; LENGTH: 4175
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(859)
US-10-017-410-1

Alignment Scores:

Pred. No.: 6e-150 Length: 4175
Score: 1404.00 Matches: 251
Percent Similarity: 96.73% Conservative: 15
Best Local Similarity: 91.27% Mismatches: 9
Query Match: 93.41% Indels: 0
DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x US-10-017-410-1 (1-4175)

```
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db ATGGGCGCCCGCACCTGGTGGGACCACTCGCGGCTGGCAGTTTCGGAGGTGGATTGGTGC 94
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db GAGGACAACTACACTATCGTCTGCTGCTATTCGAGTTCTACCAACAGATCAGCAACGTC 154
QY 41 LeuPhePheIleLeuProPheCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db TTTGTTTTTCATTTTACCTCCATCTCATGTGCTTGTTCGCCAGTACGCAACGTCGCTTC 214
QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db AACAGCGGCATCTACCTTAAATATGAGCGCTCTAGTTGTAGTGGGAGTTGATCTGTCTAC 274
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db TTTCATGCAACGCTGAGTTTCTGGGTGAGTCTGTGATGAACTTCCCATCTCTGGGTT 334
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db CTGATGTGTGCTTGGCCATCTGTTTCCAGGAGTATTTACCAAGATCTTTCGAGT 394
QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrCysLeuAla 140
Db GACAGGGCAGGTTCAGGCGAGTGGTGTGCTGCTGTCTGCAATTAACAACGCTTGGCG 454
QY 141 PheValLysProAlaIleAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db TTTATCAAGCCCGCCATCAACAATATTTCCCTGATGATTTCTGGGACTTCCATGCACTGC 514
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db CTGCTTGTGGCAGCTGAGAGGTGTGCAATGTGCGTGTGTTTAAAGTGGGCTCTTTC 574
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db TCTGGCCCTCTGTGGACTCTGCTCTCTCTGCTGATCAGGACCAAGCCCTTCTGTGAG 634
QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeu 220
Db CTGCTCTCTCTTTCATCTCCCTACCTGCACTGTGTGGTGGCATATTTCTCATCTGCCTT 694
QY 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db GCTTCGTACCTGGGCTGTGTGCTTGGCTTCTGCTGATGCTGCTCCTCAGAGATACCTGAG 754
QY 241 GlnGlyProValIleLysPheTrpProGlnLysTrpAlaPheIleGlyValProTyr 260
Db CAAGTCCAGTCACTGAGATTCGGCCCGCAGGAAATGGGCTTTTATTTGGTGTCCCTTAT 814
QY 261 ValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
Db GTGTCCTCTCTGTGTGCCCAACAAGATGCGCCAGTCAAGATCAAG 859
```

RESULT 5

US-10-182-447-6
; Sequence 6, Application US/10182447
; Publication No. US20030185814A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; APPLICANT: RADT, Marcus
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055US0
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-6

Alignment Scores:

Pred. No.: 1.9e-149 Length: 822
Score: 1390.50 Matches: 256
Percent Similarity: 94.55% Conservative: 4
Best Local Similarity: 93.09% Mismatches: 14
Query Match: 92.51% Indels: 1
DB: 15 Gaps: 1

US-10-017-410-4 (1-275) x US-10-182-447-6 (1-822)

```
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db ATGGGCGCCCGCACCTGGTGGGACCACTGGCAGCTGCGAGTGGCAGTTCGAGGTGGACTGGCGC 60
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db GAGGACAACTACCACTGCTGCTGTGCGCGAGTTCTATAACATGATCAGCAATGTC 120
QY 41 LeuPhePheIleLeuProPheCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db TTTATTTTTCATTTTACCGCCCATCTGATGCTTGTTCGTGAGTATGCAACATGCTTTC 180
QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValIleGlySerValTyr 80
Db AACAGCGGCATCTACTTAATCTGG--CTCTTGGTTGTAGCGGGAATTGGATCCGCTTAC 237
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db TTCCATGCAACCCCTTAGTTTCTGGGTGAGTCTGTGATGAACCTTTCGAGTCTCTTGGGTT 297
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db CTGATGTGTGCTTCGGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 357
QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db GACCAAGGTAGTGTTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 417
QY 141 PheValLysProAlaIleAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db TTTGTCAAGCTGCGCATCAACAACATCTCTGTGATGACCCCTGGGAGTTCTTTCGCTGCA 477
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db CTGCTCATCAGAGCTAAAGAGGTGTGACCAACATGCTGTGTGTGTGTGTGTGTGTGTGTGT 537
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
```

```
Db 538 TCGGGCTCTGGTGGACCTTGGCCCTGTTCTGCTGATCAGTACCGAGCTTCTCGGAG 597
Qy 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHisLeuLeuLeuLeuLeu 220
Db 598 CTGCTGTCATCTTCAACTTCCCTACCTGCACTGCACTGGTGGACATCTCTCATCTGCTT 657
Qy 221 AlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGlu 240
Db 658 GCTGCTACCTGGCTGTGTATGCTTGGCTACTTGTGATGCTGCTCAGAGATTCCTGAG 717
Qy 241 GlnGlyProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 260
Db 718 CAAGGCTCTGTATCAAAATTCGCGCCAGGAGAAATGGGCCCTTCATTTGGTGTCTAT 777
Qy 261 ValSerLeuLeuCysAlaAsnLysLysSerValLysLeuLeuLeuLeuLeuLeuLeu 275
Db 778 GTGTCCCTCTGTGTGCAACAGAAATCATCAGTCAAGACCAG 822

RESULT 6
US-09-945-527-48
; Sequence 48, Application US/09945527
; Publication No. US2003005588A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US2003005588A1el Nucleic Acid Molecules Encoding
; FILE REFERENCE: Nucleic Acid and Protein Homologs
; CURRENT FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-945-527-48

Alignment Scores:
Pred. No.: 5,22e-64 Length: 744
Score: 645.00 Matches: 120
Percent Similarity: 92.48% Conservative: 3
Best Local Similarity: 90.23% Mismatches: 8
Query Match: 42.91% Indels: 2
DB: 10 Gaps: 1

US-10-017-410-4 (1-275) x US-09-945-527-48 (1-744)
Qy 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db 277 ATGGCGCCCGCAGCTGGTGGACAGCTGCAGGCTGTAGCTCGAGGTGACTGGTGC 336
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 337 GAGGACAACTACACATCGTGCCTGCTATGCGCGAGTCTTACAACACGATCAGCAATGTC 396
Qy 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db 397 TTATTTTTCATTTTACCGCCCATCTGCATGTCTGTTTTCGTCATGATGCAATGCTTC 456
Qy 61 AsnSerAspIleTyrLeuLeuLeuLeuLeuValValValGlyLeuGlySerValTyr 80
Db 457 AACAGTGGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGAAATGGATCCGTCTAC 516
Qy 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db 517 TTCCATGCACCTTAGTTCCTTGGTTCAGATGCTTGTGATGAACTTGCACTCTTTGGGTT 576
Qy 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db 577 CTGATGTGTGCTTTGGCCATGTGTTTCCCGCAGAGGTATCTACCAAGATCTTTCCGAAT 636
Qy 121 AspArgGlyArgPheLysValValValSerValLeuSer 133
```

```
Db 637 GAC-----CGATACCCACCACTCACCATGATGCTTTCT 669
RESULT 7
US-10-182-447-4
; Sequence 4, Application US/10182447
; Publication No. US2003018581A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; APPLICANT: RADT, Marcus
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055U80
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-4

Alignment Scores:
Pred. No.: 9,39e-52 Length: 792
Score: 538.50 Matches: 98
Percent Similarity: 60.00% Conservative: 52
Best Local Similarity: 39.20% Mismatches: 99
Query Match: 35.83% Indels: 1
DB: 15 Gaps: 1

US-10-017-410-4 (1-275) x US-10-182-447-4 (1-792)
Qy 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 25 AGCTCCGAGGTGGACTGGTGTGAGAGCACTTCCAGTACTCGGAGCTGTGGCGAGTTC 84
Qy 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
Db 85 TACAACAGTTCTCCAAATATCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 144
Qy 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTyrThrLeuValVal 73
Db 145 CACCGGTATGCCAGAACGCTCCGCTACATTTACGTTGTCTGGTCTCTTCAATC 204
Qy 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 205 ATAGGCTGTCTCCATGTATTTCCACATGAGCTCAGCTTCTTCTGGCCAGCTGTGGAC 264
Qy 94 GluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyr 113
Db 265 GAGATCGCATCTCTGTGCTCTCTGGCAGTGGCTATGATGATGATGATGATGATGATGAT 324
Qy 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValValSerValLeuSer 133
Db 325 TTCCCTCTCTTCTTGGGGGGAACAGGTCCTCCGCTCATCTCCGCTGTCTTCACTACCACT 384
Qy 134 AlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThr 153
Db 385 GTGGTCAGACCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 444
Qy 154 LeuGlyValProCysThrAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 173
Db 445 ATTGCGCTGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 504
Qy 174 ValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIle 193
Db 505 CTTGCGCACCTGATTTGAGGTCTCCGCTGTTTATGGGCTGTGTGCTGTGCTGCTGCTG 564
```

Qy	194	SerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyrLeuHisCysMet	213
Db	565	AGTGACCCGTCTGCTTTGACGCTTCTGGCAGAGGATTCATTTCTCTATCTGCACAGATC	624
Qy	214	TrpHisIleLeuIleCysLeuAlaIaTyrLeuGlyCysValCysPheAlaTyrPheAsp	233
Db	625	TGGCATGTGCTCATCAGATCACCTTCCCTTATGGCATGTGTCAACCATGTCCTTGGTGGAT	684
Qy	234	AlaAlaSerGluIleProGluGlnGlyProValIleLysPheTyrProAsnGluLysTyr	253
Db	685	GCCAACTATGAGATGCCAGGTGAACCCCTCAAGTCCGCTACTGGCTCGGACAGTTGG	744
Qy	254	AlaPheIleGlyValProTyrValSerLeu	263
Db	745	CCCGGGGCTGCCCTACCTGGAAATC	771

RESULT 8

US-10-182-447-5

Sequence 5. Application IIS/10182447

Sequence 3, Application OS/1018,
Publication No. IIS20030185814A1

: GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: HOEMANN K&V

; AFFILIANT: HOFMANN, KAY
; APPLICANT: RADT, MARCUS

; APPLICANT: RADT, MAYCUS
: TITLE OF INVENTION: CERAMIC COE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

FILE REFERENCE: P68055US0

; CURRENT APPLICATION NUMBER: US/10/182,447

; CURRENT FILING DATE: 2002-07-29

; PRIOR APPLICATION NUMBER: PCT/EP01/00900

PRIOR FILING DATE: 2001-01-27

; PRIOR APPLICATION NUMBER: DE 10003293.1

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: DE 10011392.3

; PRIOR FILING DATE: 2000-03-09

; NUMBER OF SEQ ID NOS: 14

```

; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 5

; LENGTH: 801

; TYPE: DNA

Alignment Scores:

Argument Scores:		
Pred. No.:	8.14e-15	Length: 801
Score:	216.50	Matches: 79
Percent Similarity:	40.45%	Conservative: 29
Best Local Similarity:	29.59%	Mismatches: 105
Query Match:	14.40%	Indels: 56
DB:	15	Gaps: 9

US-10-017-410-4 (1-275) x US-10-182-447-5 (1-801)

QY	14	SerSerGluValAspTrpCysGluAsp	AsnTyrThrIleValProAlaIleA	GluPhe	33
		::::	::::	::::	
DB	43	ACCTCCACCGCTGGATGGTGGAG	GAACATCTACCTCGGTACATCGC	CGGATTC	102
		::::	::::	::::	
QY	34	TyrAsnThrIleSerAsnValLeuPhe	PheIleLeuProPrlIeCysMetCys	LeuPhe	53
		::::	::::	::::	
DB	103	TGGAATACAGTGAATAC---CTG	ATCATGNTTATACCTCCAA--TGT	TCCGGTGC	157
		::::	::::	::::	
QY	54	AspGluTyrAlaThrCysLeuAsnSer	AspIleTyrLeuIleTrpThrLeuLeu	ValVal	73
		::::	::::	::::	
DB	158	AGAGTGTGTAGACCGCTCGGAAA	AGCGGTACATGTGTTCT-TATT	TAGCACTCACAGTG	216
		::::	::::	::::	
QY	74	ValGlyIleGlySerValTyr-PheHis	PheThrLeuSerPheLeuGlyGlnMet	LeuAsp	93
		::::	::::	::::	
DB	217	GTAGAAATGGGATCCTGGTGCTT	CCACATGACTCGAAATATGAAAT	CGACTATTGGAT	276
		::::	::::	::::	
QY	94	GluLeuAlaValLeuTrpValLeuMet	CysAlaLeuAlaMetTrpPheProArg	Tyr	113
		::::	::::	::::	
DB	277	GAACTCCCAATGATATAC---AGC	TGTTCATTTGTGTACGCAATGTTT	GAAATGTTTC	333
		::::	::::	::::	
QY	114	LeuProLysIlePheArgAsnAspArg	GlyArgPheLysValValSerValLeu	Ser	133


```
US-10-017-410-4 (1-275) x US-10-302-172-875 (1-1063)
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32295
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73286E07_FLI
US-10-425-114-32295

Alignment Scores:
Pred. No.: 2,07e-13 Length: 1194
Score: 206.50 Matches: 72
Percent Similarity: 43.68% Conservative: 42
Best Local Similarity: 27.59% Mismatches: 124
Query Match: 13.74% Indels: 23
DB: 13 Gaps: 9

US-10-017-410-4 (1-275) x US-10-425-114-32295 (1-1194)
QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 117 ACATCAACGACTGAGTTGTGGAGAGAGATTATGCACACTCGTCATATATCGCAGAAATTC 176
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCysLeuPhe 53
Db 177 TACAATACCATCTCTAATGTCCCATGCGTTCTTTTGGCACTTATT-----GGATTAGTG 230
QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrIleLeuIleTrpThrLeuValVal 73
Db 231 AATGCTTTTCGCAAGGTTTGTGAAACGATTGTGTCTCTGCACATATCAATATGATA 290
QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 291 CTTCCTATTGGGAGTATGATTTCCTCCACCTTGCACTTCTCTACACACAGAGTAT 350
QY 94 GluLeuAlaValLeuTrp---ValLeuMetCysAlaLeuAlaMetTrpPheProArgArg 112
Db 351 GAGACTCTATGGTTGGAGATCTTCTCTACATGTATGTCTCTATTCACCGGACTGG 410
QY 113 TyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeu 132
Db 411 CAC-----TACAGAGACACAATGCCAACTTTCCTGTCTCTGTATGGTGCTGCC 458
QY 133 SerAlaValThrThrCysLeuAlaPheValLysProAla-----IleAsnAsnIleSer 150
Db 459 TTTCGCGTAGTACATTTCTTTGCCCGGTTCACAGTCTGATTCAAGCTGCAATACGTTGCC 518
QY 151 LeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAsp 170
Db 519 CTCTGCTTGTCTCTGCATCCCCCGGATGTACAAAGTACTACATACAGACGAAAGACGTGGGT 578
QY 171 AsnMetArgValPheLysLeuGlyLysPheSerGlyLeuTrpTrpThrLeuAlaLeuPhe 190
Db 579 GCGAAGCGGCTCGCAAAACTGTGGTC-----CTTACCTGACCCCTTGGGACCCTC 629
QY 191 CysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPhe---ProTyr 209
Db 630 TGCTGGCTGGTTGATCGGCTCTCTGCAAGAAGCTTTTCGCATTTGGTACGTCACACCGCAG 689
QY 210 LeuHisCysMetTrpHisIleLeuLeuLeuCysLeuAlaIleTyrLeuGlyCysValCysPhe 229
Db 690 GGGCAGCGGTGGTGGCAGCTGTATGGCGCTCAACTCGTGTACTATGCAAAACAG---TTC 746
QY 230 AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpPro 249
Db 747 CTGATGTTCTGCCGGGCTCAGCAG-----CGCGGTGGGAG 782
QY 250 AsnGluLysTrpAlaPheIleGly---ValProTyrValSerLeuLeuCysAlaAsnLys 268
Db 783 CGCGGATCAGCAGCTCTCTGGATTCTTGGCGTATGTTCGGTATGTCCAGGTCCAGAAACAGAGAAG 842
QY 269 Lys 269
```

```
US-10-017-410-4 (1-275) x US-10-302-172-875 (1-1063)
QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 44 ACCTCCACGCTGACTGGTGGAGGAGAACTACTCGGTGACCTCGGTGACATCGCGGAGTTTC 103
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProIleCysMetCysLeuPhe 53
Db 104 TGAATACAGTGTAGTAAAC---CTGATCATGATTATACCTCCAA--TGTTCCGTGCAATTC 158
QY 54 AspGluTyrAlaThrCysLeuAsnSerAspIleTyrIleLeuIleTrpThrLeuValVal 73
Db 159 AGAGTGTAGAGCGGTCTGGAAGCGGTACATTTGCTTCT-TATTTAGCACTCACAGTG 217
QY 74 ValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 218 GTAGGAATGGATCCCTGGTGTCTTCCATGACTCTGAAATATGAAATGCAAGTATTGGAT 277
QY 94 GluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyr 113
Db 278 GAACCTCCCAATGATATAC---AGCTGTGTCATATTGTGTACTGTCATGTTGATGTTTC 334
QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysValValValSerValLeuSer 133
Db 335 AAGATCAAGAACTCAGTAAACTACCATCTGCTTTTACCTTAGTCTTATTCAGTTTAATA 394
QY 134 AlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThr 153
Db 395 GTAACACACAGTTTACCTTAAGGTAAAGAGCCAAATATTCATCAGGTCATGTAT----- 448
QY 154 LeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArg 173
Db 449 -----GGATGTTG 457
QY 174 ValPheLysLeuGlyLeuPheSer----- 181
Db 458 GTCTTACATTAGTACCTGATCTATTATATTGTTATGTTTATCCATGGCTTAGA 517
QY 182 GlyLeuTrpTrpThr-----LeuAlaLeuPheCysTrpIleSerAsp 195
Db 518 GGAAGTGGTTATACATCATTTGGGTATATTTTATTTGGGATTTTATTTTGGAAATATAGAT 577
QY 196 ArgAlaPheCysGluLeuLeuSerSerPheAsn----- 206
Db 578 AACATATTTGTGAGTCACTGAGCACTTCGAAGAAGGTACCACCTATCATAGGTATT 637
QY 207 PheProTyrLeuHisCysMetTrpHisIleLeuLeuCysLeuAlaAlaTyrLeuGlyCys 226
Db 638 ACCACACAATTTCAATGCAATGGTGGCATATTTTAACTGGCCTTGGTTCTCTATCTTAC--- 694
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLys 246
Db 695 ATCCTTTTTCAGTTTGTATACAAAGACACTTTACCTG---AGATATAGGCCAAAGTGAAG 751
QY 247 Phe-----TyrPro 249
Db 752 TTTCTCTTGGAAATCTGGCCA 772

RESULT 10
US-10-425-114-32295
; Sequence 32295, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
```



```
Db 482 CACGAGCGATGAGACTCCAAATGGTGGAGATTCTCTATATCTTTATGACTTTAT 541
Qy 108 -----TTPheProArgTyrLeuProLys-----IlePheArgAsnAspArgGly 123
Db 542 TCACGAGACTGGCATTACCGGAGCACTATGCTTCTTTCTATACGGTGTCTGCT 601
Qy 124 -----ArgPheLysValValSerValLeuSerAlaVal 135
Db 602 TTTCAGTAGTCATTTCTGGTGGCATTCCTCAAGTGTATTCAG----- 646
Qy 136 ThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGly 155
Db 647 -----TTGCATTACGTTGGCCTCTCTCTCTATGTC 676
Qy 156 ValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPhe 175
Db 677 ATCCCAAGGATGTACAAAGTACTACATAACAATAAGACATGGCTGCCAAGCGTCTAGCA 736
Qy 176 LysLeuGlyLeuPheSerGlyLeuTyrThrLeuAlaLeuPheCysTyrTrpIleSerAsp 195
Db 737 AAGCTGTGGGT-----CTTACGTTAGCTTGGCGACTCTTGTCTGGCTATTGAT 787
Qy 196 ArgAlaPheCysGluLeuLeuSerSerPheAsnPhe---ProTyrLeuHisCysMetTrp 214
Db 788 CGAATGTTCTGTAAAGAGCTTTCACATTGGTACGTCAACCCACAAGGGCACGCAATGGTG 847
Qy 215 HisIleLeuLeuCysLeuAlaAlaTyrThrLeuGlyCysValCysPheAlaTyrPheAspAla 234
Db 848 CATATTCTCATGGGATTAACTCATACTTTCGAAACACA---TTCTTAATGTTTGGCGA 904
Qy 235 AlaSerGluLeuProGluGlnGlyProValIleLysPheTrp---ProAsnGluLysTrp 253
Db 905 GCTCAACAG-----CGTGGTGGGAGCCCAAAATTACCCAC 940
Qy 254 AlaPheIleGlyValProTyrValSerLeuLeuCysAlaAsnLysLys 269
Db 941 CTTTTCGGGTTCTTGCTTATGTCAAGATTGAGAAACCCCAAAAGAGG 988

RESULT 13
US-10-017-161-2429
; Sequence 2429, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2429
; LENGTH: 35425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(35425)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(293)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21253)..(21367)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21462)..(21603)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (23918)..(24055)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26460)..(26597)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26868)..(27016)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35125)..(35225)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (35270)..(35425)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-2429

Alignment Scores: 7.12e-09 Length: 35425
Pred. No.: 186.00 Matches: 43
Score: 44.83% Conservative: 22
Percent Similarity: 29.66% Mismatches: 35
Best Local Similarity: 12.38% Indels: 45
Query Match: 15 Gaps: 3
DB:

US-10-017-410-4 (1-275) x US-10-017-161-2429 (1-35425)

Qy 19 TrpCysGluAspAsnTyrThrIleValPro-----Ala 29
Db 21184 TGGACAGAGGCGAGCCTCTCCGCTCTCCCTATCTGACGAGCTCCTCCCTATGGCTCTCT 21243
Qy 30 IleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCys 49
Db 21244 CTGCTCCAGTTC-----TCCAATATCCCTTCTTTCATCTTCGGGCGCACTGATG 21291
Qy 50 MetCysLeuPheAspGluTyrAlaIleThrCysLeuAsnSerAspIleTyrLeuIleTyrThr 69
Db 21292 ATGCTCTCTGATGCACCGCTATGCCAGAGCGCTCCCGCTACATTTAGTTGCTGGGTC 21351
Qy 70 LeuLeuValValVal----- 74
Db 21352 CTCTTCATGATCATAGTAGGAGGTGTGTTCAGTCTGTGACAGTCGGGAGGAGGAGTGG 21411
Qy 75 -----GlyIleGlySe 78
Db 21412 GGGTTTAGGAGTGGCGGAGCCCACTGACCGCTGCCCTTCCCGCTGCGAGCGCTGTCTC 21471
Qy 78 rValTyrPheHisPheThrIleuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValIle 98
Db 21472 CATGTATTTCCACATGACGCTCAGCTTCTCTGGCCAGCTGTCTGGACGAGATCGCATCT 21531
Qy 98 uTIPValLeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePh 118
Db 21532 GTGGCTCTGGGAGTGGCTATGACATATGATGATGATGATGATGATGATGATGATGATGAT 21591
Qy 118 eArgAsnAspArg 122
Db 21592 TGGGGGGAACAGG 21604

RESULT 14
US-10-292-798-2069
; Sequence 2069, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
```



```
QY 142 ---ValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
643 GGTATTGGCTTCAAAGTGCATTATATCATCTCTCTCGGTTCCGAGAAATGTAC 702
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
703 AAGTATTACATTACACACAAAGATGTTTCAGCCAGCGGTTGCAAGCTATTTTAGTT 762
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
763 ACTTTGTATTAGAAAGTTTGTGGTTCTGT-----GATCGTGTGTTTCTGCAA 813
QY 201 LeuLeuSerSerPheAsnPhe---ProTyrLeuHisCysMetTrpHisIleLeuIleCys 219
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
814 GAGATTTCCTGGTGGCCTATTAAACCTCAGGTCATGCTTTGGCATGTTCATGGGT 873
QY 220 LeuAlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIlePro 239
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
874 TTTAATTCCTACTTTGCCAACACA---TTCTTGATGTTTTCGGGGCTCAACAG----- 924
QY 240 GluGlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValPro 259
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
925 -----CGTGGTGGTCTCCAAAAGTTCTTCATTAAATGGGGTTACCG 966
QY 260 TyrValSerLeu-----LeuCysAlaAsnLysLysSerSer 271
Db   ||||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
967 TATGTCAGATTGAGAAACCAAAAGCCAGTGATAAAAATGTGTGGAAAGAGGAAACC 1026
QY 272 Val 272
Db   :: ::
1027 CTC 1029
```

Search completed: September 18, 2004, 07:06:12
Job time : 467 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 05:11:35 ; Search time 2465 seconds
(without alignments)
3331.481 Million cell updates/sec

Title: US-10-017-410-4
Perfect score: 1503
Sequence: 1 MGRPHWDQLQASSEVDWC.....IGVPYVSLCANKSSWKIT 275

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2 1/USPTO.spool_p/US10017410/runat_15092004_164709_4703/app_query.fasta_1.455
-DB=EST -QFMT=FASTAP -SUFFIX=rbt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410@cgn 1_1 5180 @runat_15092004_164709_4703 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gaai:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1298	86.4	1022	13	BUS11164	BUS11164 AGENCOURT
2	1202	80.0	681	29	AY401889	AY401889 Homo sapi
3	1194	79.4	681	29	AY401890	AY401890 Pan trogl
4	1134	75.4	681	29	AY401891	AY401891 Mus muscu
5	1092	72.7	868	14	CA976684	CA976684 AGENCOURT
6	1092	72.7	1173	11	AK085306	AK085306 Mus muscu
7	1071.5	71.3	797	14	CA463234	CA463234 AGENCOURT
8	995	66.2	698	13	BU234223	BU234223 603792086
9	965	64.2	565	14	CF115220	CF115220 Shultzomi
10	958	63.7	781	13	BU232554	BU232554 603409105
11	922	61.3	622	14	CB723138	CB723138 UI-M-GHO-
12	916	60.9	633	10	BB660847	BB660847 BB660847
13	906	60.3	648	12	BU606795	BU606795 BJ060795
14	896	59.6	847	13	BU220481	BU220481 603107516
15	882	58.7	937	13	EX842808	EX842808 EX842808
16	856	57.0	678	13	BU203269	BU203269 604153603
17	850	56.6	548	14	CF169369	CF169369 B0812607-
18	811	54.0	498	12	BI848265	BI848265 470659 MA
19	808	53.8	449	13	EX646596	EX646596 DXEPD781B
20	715	47.6	432	14	CF169808	CF169808 B0818008-
21	682.5	45.4	666	13	BW215750	BW215750 BW215750
22	680	45.2	803	14	CF593785	CF593785 AGENCOURT
23	673	44.8	401	10	BF554219	BF554219 UI-R-CO-h
24	672	44.7	454	14	CB784796	CB784796 AMGNNUC:T
25	660.5	43.9	815	13	EX623644	EX623644 BX623644
26	636.5	42.3	741	13	BW216493	BW216493 BW216493
27	634	42.2	434	14	CB758986	CB758986 AMGNNUC:M
28	623	41.5	427	9	AA900336	AA900336 UI-R-EO-c
29	622	41.4	399	14	CB706492	CB706492 AMGNNUC:M
30	609.5	40.6	480	14	CB728612	CB728612 AMGNNUC:M
31	609	40.5	976	14	CA967276	CA967276 CCLX06a22
32	601	40.0	485	10	BF549345	BF549345 UI-R-AO-a
33	593	39.5	789	14	CA969385	CA969385 CCLX06a22
34	578.5	38.5	712	13	BX620332	BX620332 BX620332
35	558	37.1	723	12	BM602328	BM602328 170006870
36	535.5	35.6	398	14	CB707581	CB707581 AMGNNUC:M
37	528	35.1	634	9	AL864302	AL864302 AL864302
38	524.5	34.9	806	14	CD052625	CD052625 LITHZF000
39	518	34.5	626	13	BW220052	BW220052 BW220052
40	518	34.5	715	13	BU337163	BU337163 603514014
41	517.5	34.4	1562	11	AK075884	AK075884 Mus muscu
42	517.5	34.4	2429	11	AK028901	AK028901 Mus muscu
43	513	34.1	685	12	BJ062108	BJ062108 BJ062108
44	510	33.9	357	13	BY168309	BY168309 BY168309
45	499	33.2	295	10	BE668106	BE668106 156540 MA

ALIGNMENTS

RESULT 1
BUS11164
LOCUS BUS11164 1022 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT 10107530 NIH MGC 134 Mus musculus cDNA clone
IMAGE:6505924 5', mRNA sequence.
ACCESSION BUS11164
VERSION BUS11164.1 GI:22817397
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1022)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LNA4067 row: j column: 05
 High quality sequence stop: 681.
 Location/Qualifiers

FEATURES

source

```
1..1022
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:10090"
    /clone="IMAGE:6505924"
    /tissue_type="undifferentiated limb"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH MGC 134"
    /notes="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
    Cloned unidirectionally. Primer: Oligo dt. Average insert
    size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
    this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 6,01e-128 Length: 1022
Score: 1298.00 Matches: 244
Percent Similarity: 92.86% Conservative: 16
Best Local Similarity: 87.14% Mismatches: 14
Query Match: 86.36% Indels: 6
DB: 13 Gaps: 0
```

US-10-017-410-4 (1-275) x BU511164 (1-1022)

```
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db 120 ATGGGCGCCCGCAGCTGGTGGGACCCCTGGGCTGGGAGTGGAGTGGTGGTGC 179
QY 21 GluAspAsnTrpThrLeuValProAlaIleAlaGluPheTrpAsnThrLeuSerAsnVal 40
Db 180 GAGGACAACTACACTATCGTCCCTGGCCTTGGCAGTCTTACAAACAGATCAGCAAGTC 239
QY 41 LeuPhePheIleLeuProPheCysMetCysLeuPheAspGluTrpAlaThrCysLeu 60
Db 240 TTGTTTTTCATTTTACTCCCATCTCCATCTCATGTGCTTGTTCGCCAGTACGCAAGTCTTC 299
QY 61 AsnSerAspIlePyrLeuIleTrpThrLeuValValValValGlyIleGlySerValTrp 80
Db 300 AACAGCGGCATCTACTTATATGAGCGCTCCTAGTTGTAGTGGGATGGATCTGTCTAC 359
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db 360 TTCCATGCAACGCTGAGTTTCTGGTGCAGATCTTGTATGATGAATTCCTTCCTGGT 419
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTrpLeuProLysIlePheArgAsn 120
Db 420 CTGATGTGTGCTTGGCCATGTGTTTCCAGAGGATATTTACCAAGATCTTTCGGAT 479
QY 121 AspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db 480 GACAGGGCAGGTTCAAGGCGAGTGGTGTGCTCTGTCTGCAATTACACAGTCTGGCG 539
QY 141 PheValIleProAlaIleAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db 540 TTATCAAGCCGCCATCAACATATTTCCCTGATGATTCTGGGACTTCCATGCACTGGG 599
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
```

```
Db 600 CTGCTTGTGTCAGAGCTGAAGAGGTGTGACAAATGTGCGTGTGTTTAAAGTGGGCTCTTC 659
QY 181 SerGlyLeuTrpTrpThrLeu-AlaLeuPheCysTrpIleSerAspArgAlaPheCysG 200
Db 660 TCTGGCCTCTGGTGGACTCTGGGCTCTCTTCTGCTGGATCAGCAGCAAGCCTTCTGTGA 719
QY 200 uLeuLeuSerSerPheAsnPheProTrpLeuHisCysMetTrpHisIleLeuIleCysLe 220
Db 720 GCTGCTCTCTCTTTCACCTTCCCTACCTGCACTGTGTGGGATATTTCTATCTGCT 779
QY 220 uAlaAlaTrpLeuGlyCysValCysPheAlaTrpPheAspAlaAlaSerGluIleProG 240
Db 780 TGCTTCGTACCTGGGCTGTGTGCTTCCCTACTTTGATGCTGCTCAGAGATACCTGA 839
QY 240 uGlnGlyProVal-IleLysPheTrp-ProAsnGluLysTrpAlaPhe-IleGlyValP 259
Db 840 GCCAAGTCCAGTCCATCAGATTCTGGGCCACAGAAAGTGCAGTCAAGATC 899
QY 259 roTrpValSerLeuLeuCysAla-AsnLysLysSerSerValLysIle 274
Db 900 CCTATGGTCCCTTCTGTGTGCCACAGAAAGTGCAGTCAAGATC 947
```

RESULT 2

AY401889

LOCUS

DEFINITION

Homo sapiens

AY401889

VERSION

AY401889.1

KEYWORDS

GSS.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-017-410-4 (1-275) x AY401889 (1-681)

QY

50 MetCysLeuPheAspGluTrpAlaThrCysLeuAsnSerAspIleTrpLeuIleTrpThr 69

Db	1	ATGTGCTTGTTGGTCAGTATGCAACATGCTTCAACAGATGGCATCTACTTAATCTGGACT	60
Qy	70	LeuLeuValValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGly	89
Db	61	CTTTTGGTGTAGTGGGAAATGGATCCGTCATCTCCATTTTACCCCTTAGTTCTTGGGT	120
Qy	90	GlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe	109
Db	121	CAGATGCTTCATGACCTTGCAGTCCCTTTGGTTCGTATGTGTGCTTTGGCCATGTGTTTC	180
Qy	110	ProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysValValVal	129
Db	181	CCCAGAAGGTATCTACCAAAAGATCTTTCCGAATGACCCGGGTAGGTTCAAGGTGGTGTC	240
Qy	130	SerValLeuSerAlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIle	149
Db	241	AGTGTCTGTCTGGGTTTACGACGTGCTGGCATTGTCAAGCCTGCCATCAACACATC	300
Qy	150	SerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCys	169
Db	301	TCTCTGATGACCCCTGGGAGTTCCCTTGCACTGCATCTCATCGCAGAGCTAAGAGGTGT	360
Qy	170	AspAsnVetArgValPheLysLeuGlyLeuPheSerGlyLeuTyrTrpThrLeuAlaLeu	189
Db	361	GACAACATGCGTGTGTGTTTAAGCTGGGCTCTTCTCGGSCCTCTCGTGAGACCTGGCCCTG	420
Qy	190	PheCysThrTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyr	209
Db	421	TTCTGCTGGATCAGTACCGAGCTTTCTGCGAGCTGCTGTCTATCCTTCAACTCCCTTAC	480
Qy	210	LeuHisCysMetTrpHisIleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPhe	229
Db	481	CTGCACATGCATGTGGCACATCCTCATCTGCCTTGTGCTGCCTACCTGGGTGTGTATGCTTT	540
Qy	230	AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpPro	249
Db	541	GCCTACTTTTGTATGCTGCCCTCAGAGATTCTGAGCAAGGCCCTGTCTCATCAAGTTCTGGGCC	600
Qy	250	AsnGluLysTyrAlaPheIleGlyValProTyrValSerLeuLeuCysAlaAsnLysLys	269
Db	601	AATGAGAAATGGGCCCTTCATTGGTGTCCCTTATGTGCTCTCTGTGTGCCAACAGAA	660
Qy	270	SerSerValLysIleThr	275
Db	661	TCATCAGTCAAGATCAAG	678

RESULT 3			
AY401890			GSS 12-DEC-2003
LOCUS	681 bp	DNA	linear
DEFINITION	pan troglodytes HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		

KEYWORD	SS.
SOURCE	Pan troglodytes (chimpanzee)
ORGANISM	Pan troglodytes
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. 1 (bases 1 to 681)

TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
 source 1..681 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..>681
 /locus_tag="HCM1041"
ORIGIN
 Alignment Scores:
 Pred. No.: 4,51e-117 Length: 681
 Score: 1194.00 Matches: 221
 Percent Similarity: 98.23% Conservative: 1
 Best Local Similarity: 97.79% Mismatches: 4
 Query Match: 79.44% Indels: 0
 DB: 29 Gaps: 0
 US-10-017-410-4 (1-275) x AY401890 (1-681)
 Qy 50 MetCysLeuPheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTrpThr 69
 Db 1 ATGTGCTTGTTTCGTAGTAGTCACACATGCTTCAACAGTGGCATCTACTTAATCTGGACT 60
 Qy 70 LeuLeuValValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGly 89
 Db 61 CTTTGTGTTGTATGFGGGAATTGGATCCGTCTACTTCCATGCCAACCTTAGTTCTTGGGT 120
 Qy 90 GlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe 109
 Db 121 CAGATGCTTGATGAACCTTGCACTCCTTTGGTTCCTGATGTGTGCTTGGCCATGTGTC 180
 Qy 110 ProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysValValVal 129
 Db 181 CCCAGAAGGTATCTACCAAAAGATCTTTCGGAATGCCGGGTAGGTTCAAGGTGGTGTC 240
 Qy 130 SerValLeuSerAlaValThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIle 149
 Db 241 AGTGTCTGTCTGGGGTTTACGACATGCTTGGCAATTTGCAAGCTGCCATCAACACATC 300
 Qy 150 SerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCys 169
 Db 301 TCTCTGATGACCTGGGAGTTCCTTGCACTGCACCTGCTCATTTGCAGAGCTAAAGAGGTG 360
 Qy 170 AspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeu 189
 Db 361 GACAAACATGCGTGTGTTTAAAGCTGGGTCTCTTCTCGGGCTCTGTGGGACCTGGCCCTG 420
 Qy 190 PheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyr 209
 Db 421 TTCTGCTGGATCAGPGACCGAGCTTCTCGGAGCTGTGTCTATCTTCACTTCCCTTAC 480
 Qy 210 LeuHisCysMetTrpPheIleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPhe 229
 Db 481 CTGCACTGATGTGGCAATCTCATCTGCTTGGCTTGTGCTACCTGGCTGTGTATGCTTT 540
 Qy 230 AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpPro 249
 Db 541 GCCTACTTTGATGCTGCTCAGAGATTCCTGAGCAAGGCCCTGTCTCATCAAGTTCTGGCCC 600
 Qy 250 AsnGluLysTrpAlaPheIleGlyValProTyrValSerLeuLeuCysAlaLeuLysLys 269
 Db 601 AATGAGAAATGGGCTTCATTGGTGTCCCTATGTCTCCTCTGTGTGCCAATAAGAAA 660
 Qy 270 SerSerValLysIleThr 275
 Db 661 TCATCAGTCAAGATCACG 678
RESULT 4

```

AY401891      681 bp      DNA      linear      GSS 12-DEC-2003
LOCUS      Mus musculus HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY401891
VERSION      AY401891.1 GI:39757877
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 681)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source      Location/Qualifiers
1..681
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    <1..>681
    /locus_tag="HCM1041"
ORIGIN
Alignment Scores:
Pred. No.:      1.16e-110      Length:      681
Score:      1134.00      Matches:      204
Percent Similarity:      96.46%      Conservative:      14
Best Local Similarity:      90.27%      Mismatches:      8
Query Match:      75.45%      Gaps:      0
DB:      29
US-10-017-410-4 (1-275) x AY401891 (1-681)
QY      50 MetCysLeuPheAspGluTyrAlaThrCysLeuAsnSerAspIleTyrLeuIleTrrpThr 69
Db      1 ATGTGCTTTGTCGCCAGTACGACGCTGCTTCAACAGCGGCATCTACTTAATATGGAGC 60
QY      70 LeuLeuValValGlyIleGlySerValTyrPheHisPheThrLeuSerPheLeuGly 89
Db      61 CTCCTAGTTGTAGTGGGATTGATGCTCTACTTCCATGCAACGCTGAGTTCCTGGGT 120
QY      90 GlnMetLeuAspGluLeuAlaValLeuTrrpValLeuMetCysAlaLeuAlaMetTrrpPhe 109
Db      121 CAGATGCTTGATGAACATTGCCATCTGTGGTCTCTGATGTGCTTTGGCCATGGTGT 180
QY      110 ProArgArgTyrLeuProIlystlePheArgAsnAspArgGlyArgPheIlystleValVal 129
Db      181 CCCAGAGGTATTACCAAGATCTTTCGGAATGACAGGGCAGGTTCAGGCGAGTGGTG 240
QY      130 SerValLeuSerAlaValThrThrCysLeuAlaPheValIlystProAlaIleAsnAnile 149
Db      241 TGTGTCTCTGTGCAATTACAGCTGTGGCTTTCATCAAGCCGCCCATCAACATATT 300
QY      150 SerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuIlystArgCys 169
Db      301 TCCCTGATGATCTGGGACTTCCATGCATCGCTGCTTGTTCAGAGAGTGAAGAGGTGT 360
QY      170 AspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeuTrrpTrrpThrLeuAlaLeu 189
Db      361 GACAATGTCGTGTGTTTAAGCTGGGCTCTTCTCTGGCTCTGGTGGACTCTGGCTTC 420
QY      190 PheCysTrrpIleSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTyr 209
Db      421 TTCTGCTGGATCAGCGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCT 480
QY      210 LeuHisCysMetTrrpHisIleLeuLeuLeuAlaPheCysLeuAlaPheCysValCysPhe 229
Db      481 CTGCACCTGTGTGGGCATATTCTCATCTGCTTGTACCTGGGCTGTGTGTGTTTC 540
QY      230 AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrrpPro 249
Db      541 GCCTACTTTGATGCTGCTCCTCAGAGATACCTGAGCAAGGTCCAGTCATCATGATTTGSCCC 600
QY      250 AsnGluLysTrrpAlaPheIleGlyValProTrrpValSerLeuLeuCysAlaAsnLysLys 269
Db      601 AGCGAGAAATGGGCTTTTATTTGGTGTCCCTTATGTCTCTCTCTCTCTCTCTCTCTCT 660
QY      270 SerSerValIlystleThr 275
Db      661 TCGCCAGTCAAGATCAGC 678
RESULT 5
CA976684      868 bp      mRNA      linear      EST 06-JAN-2003
LOCUS      AGENCOURT 8877974 NCI CGAP_Mam2 Mus musculus cDNA clone
DEFINITION      IMAGE:6437893 5', mRNA sequence.
ACCESSION      CA976684
VERSION      CA976684
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13955 row: g column: 14
High quality sequence stop: 679.
FEATURES
    Location/Qualifiers
1..868
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="FVB/N-3"
    /db_xref="taxon:10090"
    /clone="IMAGE:6437893"
    /tissue_type="tumor, biopsy sample"
    /dev_stages="5 months"
    /lab_host="DH10B"
    /clone_lib="NCI CGAP Mam2"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
    Library constructed by Life Technologies. Investigator
    providing samples: Gilbert Smith, NIH"
ORIGIN
Alignment Scores:
Pred. No.:      4.92e-106      Length:      868
Score:      1092.00      Matches:      195
Percent Similarity:      96.26%      Conservative:      11
Best Local Similarity:      91.12%      Mismatches:      8

```



```
/codon_start=1
/protein_id="BAC39416.1"
/db_xref="GI:26351559"
/translation="MGAPHWDLRAGSSEVDWCDENYTIIVPAEYNTISNVLPFI
LPPTCMCLFRQYATCFNSGVIYLITWLLVWVGISGVYFHTLSPFAGMLDELAILVLM
CALAMWPRVLPKIPENDRGRKAVCVLSAITTCIAFKPALNYSIMLIGLPCTA
LLVAELAKRCDNRVFKLGLPSGLWTLALFCWISDQAFCELLSSFFHFFYLHCWVSADR
G"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 7,36e-106 Length: 1173
Score: 1092.00 Matches: 195
Percent Similarity: 96.26% Conservative: 11
Best Local Similarity: 91.12% Mismatches: 8
Query Match: 72.65% Indels: 0
DB: 11 Gaps: 0
```

US-10-017-410-4 (1-275) x AK085306 (1-1173)

```
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db |||||
QY 72 ATGGGGCGCCCGCACCTGGTGGGACCACTCGGGCTGGAGTTCGGAGGTGGATTGGTGC 131
Db |||||
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db |||||
QY 132 GAGGACACACTACTATCGTCCCTGCCATTCGCGAGTTCACACACGATCAGCAACGTC 191
Db |||||
QY 41 LeuPhePheIleLeuProPheCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db |||||
QY 192 TTGTTTTCATTTTACCTCCCATCTGATGCTGTTTCCGCCAGTACGCAAGTCTTC 251
Db |||||
QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db |||||
QY 252 AACAGCGGCATCTACTAATATGAGCGCTCTAGTTGTAGTGGGATGGATCTGCTAC 311
Db |||||
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db |||||
QY 312 TTCCATGCAACGCTGAGTTCTCTGGGTGATGCTTGTAGTGAATTCGCCATTTCTGGGTT 371
Db |||||
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsn 120
Db |||||
QY 372 CTGATGTGTGCTTTGGCCATGCTGTTTCCAGGAGTATTTACCAAGATCTTTCGAT 431
Db |||||
QY 121 AspArgGlyArgPheIysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db |||||
QY 432 GACAGGGGAGGTTCAAGGAGTGTGTGTCTCTGCAATTACAACGCTCTTGGCG 491
Db |||||
QY 141 PheValLysProAlaIleAsnAniIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db |||||
QY 492 TTTATCAAGCCGCCCATCAACATATTTCCCTGATGATTTCTGGGACTTCCATGCACTGGC 551
Db |||||
QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db |||||
QY 552 CTGCTTGTGTGAGAGTGAAGAGTGTGACATGTGCGTGTGTAAAGTGGGCTCTTC 611
Db |||||
QY 181 SerGlyLeuTrpThrIleAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200
Db |||||
QY 612 TCTGGCCTCTGTGGACTCTGCTCTCTCTCTGCTGATCAGGAGCAACGCTCTCTGTGAG 671
Db |||||
QY 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTrp 214
Db |||||
QY 672 CTGCTCTCTCTCTTCACTTCCCTACCTACCTGCACTGTGTGTGG 713
Db |||||
```

RESULT 7

CA463294

LOCUS

DEFINITION AGNCOURT 10691314 NTH MGC 169 Mus musculus cDNA clone

IMAGE:6770649 5', mRNA sequence.

ACCESSION CA463294

VERSION CA463294.1 GI:24919646

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
(bases 1 to 797)

REFERENCE

NIH-MGC <http://mgs.nci.nih.gov/>.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jonathan Kuo, NIMH

cDNA Library Preparation: Michael Brownstein Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM3088 row: p column: 08

High quality sequence stop: 536.

FEATURES

source

1..797
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6770649"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH MGC 169"
/notes="Organ: Testicles; Vector: pDNR-LIB; Site 1: SfiI
(ggccattagccc); Site 2: SfiI (ggccctcggcc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3',
adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAGCAGATGGCATTCAGCGCGGG-3' and
5'-ATTCTAGAGCGGCGGCGGCACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

```
Pred. No.: 6.8e-104 Length: 797
Score: 1071.50 Matches: 207
Percent Similarity: 86.26% Conservative: 19
Best Local Similarity: 79.01% Mismatches: 15
Query Match: 71.29% Indels: 21
DB: 14 Gaps: 3
```

US-10-017-410-4 (1-275) x CA463294 (1-797)

```
QY 1 MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db |||||
QY 35 ATGGGGCGCCCGCACCTGGTGGGACCACTCGGGCTGGGAGTTCGGAGGTGGATTGGTGC 94
Db |||||
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db |||||
QY 95 GAGGACAACTACACTATCTGCTGCCATTCGCGAGTTCACACACGATCAGCAACGTC 154
Db |||||
QY 41 LeuPhePheIleLeuProPheCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
Db |||||
QY 155 TTGTTTTTCATTTTACCTCCCATCTGCGAGTGTGTTCCGCCAGTACGCAACGTCCTTC 214
Db |||||
QY 61 AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db |||||
QY 215 AACAGCGGCATCTACTAATATGGAGCTCTCTAGTTGTAGTGGGATGGATCTGCTAC 274
Db |||||
QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db |||||
QY 275 TTCCATGCAACGCTGAGTTTCTGCGTCCAGATGCTTGTATGAATTCGCCATTTCTGGGTT 334
Db |||||
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsn 120
Db |||||
QY 335 CTGATGTGTCTTTGGCCATGTGGTTTCCAGGAGGTATTTACCAAGATCTTTCGGAT 394
Db |||||
QY 121 AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db |||||
```

|||||GACAGGCGAGTTCAAGGAGTGGTGTGCTGCTGCTCAATTACAAGCTGCTGGCG 454
 |||||PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
 |||||TATCAAGCCCGCATCAACAATATTTCCCTGATGATTCTGGGACTTCCATGCACTGCG 514
 |||||LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
 |||||CTGCTCTTGTTCAGAGCTGAAGAGGTGACAAATGTGGTGTGTTAAAGCTGGGCTCTTC 574
 |||||SerGlyLeuIleThrLeuAlaLeuPheCysTrpIleSerAspArgAla-PheCysGly 200
 |||||TCTGGCCTCTGGTGAGTCTGGCTCTCTTCTGCTGGATCAGCAGCAGCCCTTCTGGA 634
 |||||uLeuLeuSer-SerPheAsnPheProTyrLeu-HisCys-MetTrpHisIle-----Le 217
 |||||GCTGCTCTCCCTCTTTCACCTTCCCTACCTGGGACTGTGGTGGGCAATATTCCTCT 694
 |||||uLeuLeuAlaAlaTyrLeuGlyCysValCysPheAla-----TyrPheAspAla 234
 |||||CCTGCCCTTGGCTTCCACCTGCGGCGCTGTGTGGCGCTTCCGCCCTACTTT----- 749
 |||||aAlaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrp 253
 |||||-----TTGAGGCGCTGGCCTCCAAAAATAC 776

RESULT B
 BU234223 698 bp mRNA linear EST 26-NOV-2002
 LOCUS 603792086F1 CSEQHN24 Gallus gallus cDNA clone ChEST757k21 5', mRNA
 DEFINITION sequence.

ACCESSION BU234223.1 GI:25478587
 VERSION BU234223.1
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

TITLE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 JOURNAL A Comprehensive Collection of Chicken cDNAs
 MEDLINE Curr. Biol. 12 (22), 1965-1969 (2002)
 PUBMED 22335534

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source
 1..698
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="ChEST757k21"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_lib="CSEQHN24"
 /note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores: 8,48e-96 Length: 698
 Pred. No.: 995.00 Matches: 177
 Score: 995.00 Conservative: 14
 Percent Similarity: 95.50% Mismatches: 9
 Best Local Similarity: 88.50% Indels: 0
 Query Match: 66.20% Gaps: 0
 DB: 13

US-10-017-410-4 (1-275) x BU234223 (1-698)

Qy 76 IleGlySerValTyrPheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeu 95
 Db 1 ATTCGATCTGTTTACTTTTCATGTCACCTCAGCTTCCTGGGTCTGATGCTGATGAGCTG 60
 Qy 96 AlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuPro 115
 Db 61 GCTATTCTCTGGGTCCTTATGTGTGCTCTGCCAGTGGTTCCCTAGAGATATCTACCA 120
 Qy 116 LysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeuSerAlaVal 135
 Db 121 AGAGTTTTCGAATGACAGAGCGGTTTAAAGCTGCTGTGGTCTCTCTGGAGTT 180
 Qy 136 ThrThrCysLeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGly 155
 Db 181 ACTACCTGCTTGGCTTCATTAAACCTGCCATCAACAACATCTCACTAATGACTCTGGT 240
 Qy 156 ValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPhe 175
 Db 241 GTTCTTCACAGCTTTTACTCATTCGTGAGTGAAGAGGTGTGAACCTCTGGTGTGTAC 300
 Qy 176 LysLeuGlyLeuPheSerGlyLeuTrpThrLeuAlaLeuPheCysTrpIleSerAsp 195
 Db 301 AAGCTTGTCTGTCTTTCAGGCTCTTGGTGGATGTAGCAGCTTTCTGCTGGATCAGTAC 360
 Qy 196 ArgAlaPheCysGluLeuSerSerPheAsnPheProTyrLeuHisCysMetTrpHis 215
 Db 361 AAAGCTTTTGTGAGATCTGCTCATCATTAACCTCCCTATTTCCCTATTTGCACTGTGTAGGC 420
 Qy 216 IleLeuIleCysLeuAlaAlaTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAla 235
 Db 421 ATTTTGATTTGCTTGGCGCTTACCTAGGCTGTGCTGTTCGCTTACTTCGATGCTGCC 480
 Qy 236 SerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaPhe 255
 Db 481 TCCGAGATCCCTGAGCAGGCGCCCTCATAAAGTTCTGGCCAAGTGAAGATGGGCATTC 540
 Qy 256 IleGlyValProTyrValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleThr 275
 Db 541 ATTGGCGTTCCCTACGCTCACCTCTCTGTGTGCACACAAGAAATCACCGGTGAAGATCACA 600

RESULT 9
 CF115220 565 bp mRNA linear EST 23-JUL-2003
 LOCUS Shultzomica08471 Rat lung airway and parenchyma cDNA libraries
 DEFINITION Rattus norvegicus cDNA clone NP6159 5', mRNA sequence.

ACCESSION CF115220
 VERSION CF115220.1 GI:33175919
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 565)
 Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,

Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
 Plummer, C.G. and Buckpitt, A.R.
 Gene Expression Analysis in Response to Lung Toxicants: I.
 Sequencing and Microarray Development
 Unpublished (2003)
 Contact: Shultz, MA
 Dept. of Molecular Biosciences, School of Veterinary Medicine
 University of California, Davis
 1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
 Tel: 530 752 0793
 Fax: 530 752 4698
 Email: mashultz@ucdavis.edu
 Average Phred score is 20 or better. All poor quality data (Phred <
 20) and vector/linker sequence has been removed.
 High quality sequence stop: 565.

FEATURES

Location/Qualifiers
 1..565
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="NP6159"
 /sex="male"
 /tissue_type="airway or parenchyma"
 /dev_stage="adult"
 /clone_lib="Rat lung airway and parenchyma cDNA libraries"
 /note="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
 Site 2: Not I; mRNA was isolated from microdissected rat
 lung airways and parenchyma tissues."

ORIGIN

Alignment Scores:

Pred. No.: 1.02e-92 Length: 565
 Score: 965.00 Matches: 177
 Percent Similarity: 97.34% Conservative: 6
 Best Local Similarity: 94.15% Mismatches: 4
 Query Match: 64.20% Indels: 1
 DB: 14 Gaps: 0

US-10-017-410-4 (1-275) x CF115220 (1-565)

QY 74 ValGlyTleGlySerValTyrPheHisPhe-ThrLeuSerPheLeuGlyGlnMetIleuAs 93
 Db 3 GTGGGAATCGAATCTCTACTCTCCATNGCAACTCTTAGTTTCCTGGGTGAGATTTGA 62
 QY 93 pGluLeuAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTy 113
 Db 63 TGAACCTGCCATCTTTGGGTCTGTATGTGTCTTTGCCATGTGTGTTCCTCCAGAGTGA 122
 QY 113 rLeuProLysIlePheArgAsnAspArgGlyArgPheLysValValSerValLeuSe 133
 Db 123 TTTACCAAGATCTTTCCGAATGACAGGGCAGGTTCAAGGCAGTGTGTGTCTGTCTGTC 182
 QY 133 rAlaValThrCysLeuAlaPheValLysProAlaIleAsnIleSerLeuMetTh 153
 Db 183 TGCAATTACCAATGCTCTGGGTTTGCAAGCTGCTATCAAAATATCTCTCTGATGAT 242
 QY 153 rLeuGlyValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetAr 173
 Db 243 TCTGGGGTTCGGTCACTGGCTGCTCATTTGACAGAGCTGACAGAGTGTACATGTGCG 302
 QY 173 gValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIl 193
 Db 303 TGTGTTTAACTGGGCTCTCTCTGGGCTTTGGTGGACTCTGCTCTCTCTGCTGGAT 362
 QY 193 eSerAspArgAlaPheCysGluLeuLeuSerSerPheAsnPheProTrpLeuHisCysMe 213
 Db 363 CATGACCGAGCTCTGTGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 422
 QY 213 tTrpHisIleLeuIleCysLeuAlaLafyrLeuGlyCysValCysPheAlaTrpPheAs 233
 Db 423 GTGGCACATCTCATCTGCTTGGCGGTACCTGGGCTGCTGTCTGCTGCTGCTGCTGCTGCT 482

QY 233 pAlaAlaSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTr 253
 Db 483 TGCTGCCTCAGATACCGAAGGTCAGATCATCAGATTCGCCCCAGTGAATG 542
 QY 253 pAlaPheIleGlyValProTyr 260
 Db 543 GCCTTTTATCGGTGTCCTCAT 564
 RESULT 10
 LOCUS BU232554 781 bp mRNA linear EST 26-NOV-2002
 DEFINITION 603409105F1 CSEQHN24 Gallus gallus cDNA clone CHEST323j2 5', mRNA
 sequence.
 ACCESSION BU232554
 VERSION BU232554.1 GI:25475352
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 781)
 AUTHORS Boardman, P.E., Sanz-Esquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..781
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST323j2"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_lib="CSEQHN24"
 /note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 8.84e-92 Length: 781
 Score: 958.00 Matches: 176
 Percent Similarity: 94.06% Conservative: 14
 Best Local Similarity: 87.13% Mismatches: 10
 Query Match: 63.74% Indels: 2
 DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x BU232554 (1-781)

QY 76 IleGlySerValTyrPheHis-PheThrLeuSerPheLeuGlyGlnMetLeuAspGluLe 95

```

Db      1  ATTGGAATCTGTTACTTTCATGGCCACCTCAGCTTCTGGGTGAGATGCTGGATGAGCT 60
Qy      95  uAlaValLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTrpLeuPr 115
Db      61  GCCTATTCTCTGGTCTTATGTGCTCTTCCCATGGTGGTTCCTAGGATATCTACC 120
Qy      115  oLysIlePheArgAsnAspArgGlyAcqPheLysValValSerValLeuSerAlaVal 135
Db      121  AAGAGTCTTTTGAATGACAGAGCGGTTTAAAGCTGCTGTTGGTGTCTTCTGAGT 180
Qy      135  lThrThrCys-LeuAlaPheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuG 155
Db      181  TACTACCTGGCTTGGCTTCATTAAACCTGCCATCAACACATCTCACTAATGACTCTGG 240
Qy      155  lValProCysThrAlaLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValP 175
Db      241  GTGTTCTCTGCACAGCTTTACTCATTTGCTGAGTTGAAGAGGTGTGAACACCTGGCTGT 300
Qy      175  heLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerA 195
Db      301  ACAAGCTTGGTCTGTTTCAGGCTCTTTGGTGATGCTAGCACTTTTCTGCTGGATCATG 360
Qy      195  sPArgAlaPheCysGluLeuSerSerPheAsnPheProTrpTrpLeuHisCysMetTrpH 215
Db      361  ACAAGCTTTTGTGAGATCTGGTCATCATTTAACTTCCCTATTTCACCTGTGTATGGC 420
Qy      215  iIleLeuIleCysLeuAlaLysLeuGlyCysValCysPheAlaLysPheAspAlaA 235
Db      421  ACATTTGATTTGCTTGGCTTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy      235  laSerGluIleProGluGlnGlyProValIleLysPheTrpProAsnGluLysTrpAlaP 255
Db      481  CTTCCGAGATCCTCAGACAGGCGCCGTCATTAAGTTCTGGCAAGTAGAGATGGGAT 540
Qy      255  heIleGlyValProTrpValSerLeuLeuCysAlaAsnLysLysSerSerValLysIleT 275
Db      541  TCATTTGGGCTTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy      275  hr 275
Db      601  CA 602

CB723138      622 bp      mRNA      linear      EST 09-JUL-2003
UI-W-GHO-ceh-f-04-0-UI.r1 NIH BMAP_GHO Mus musculus cDNA clone
IMAGE:6839525 5', mRNA sequence.
CB723138
CB723138.1  GI:29780280
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 622)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .622
/organism="Mus musculus"

```

FEATURES

source

```

/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6839525"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 4,58e-88      Length: 622
Score: 922.00      Matches: 167
Percent Similarity: 76.75%      Conservative: 8
Best Local Similarity: 73.25%      Mismatches: 7
Query Match: 61.34%      Indels: 46
DB: 14      Gaps: 1

```

US-10-017-410-4 (1-275) x CB723138 (1-622)

```

Qy      1  MetGlyAlaProHisTrpTrpAspGlnLeuGlnAlaGlySerSerGluValAspTrpCys 20
Db      76  ATGGCGCGCCGCACTGGTGGGACCACCTGGCGGCTGGCAGTTCGGAGGTGGATTTGGTGC 135
Qy      21  GluAspAsnTrpTrpIleValProAlaIleAlaGluPheTrpAsnThrIleSerAsnVal 40
Db      136  GAGGAACTACATCATCTGCTGCCATTCGCCAGTTCCTACACACGATCAGCAGCTC 195
Qy      41  LeuPhePheIleLeuProIleCysMetCysLeuPheAspGluTrpAlaThrCysLeu 60
Db      196  TTGTTTTCATTTTACCTCCCATCTGCATGTGCTTTTCCGCGCAGTACGCAACGTCTTC 255
Qy      61  AsnSerAspIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db      256  AACAGCGGCATCTACTTAATATGAGCGCTCTCTAGTTAGTGGGATTTGGATCTGTCTAC 315
Qy      81  PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpVal 100
Db      316  TTCCATGCAACGCTGAGTTTCTCGGTGAGTCTGTATGAACTTGCCATCTGTGGGTT 375
Qy      101  LeuMetCysAlaLeuAlaMetTrpPheProArgArgTrpLeuProLysIlePheArgAsn 120
Db      376  CTGATGTGCTTTGGCCATGTGGTTTCCAGAGGTATTTACCAAGATCTTGGGAAT 435
Qy      121  AspArgGlyArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAla 140
Db      436  GAC-----AGGTGTGACATGTGCGTGTGTTTAAAGCTGGGCTCTTC 438
Qy      141  PheValLysProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
Db      438  ----- 438
Qy      161  LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPhe 180
Db      439  -----AGGTGTGACATGTGCGTGTGTTTAAAGCTGGGCTCTTC 477
Qy      181  SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGlu 200

```

478 TCTGGCCTCTGGTGGACTCTGGCTCTCTCTCTGGTGGATCAGCAGCCAGCCCTCTGTGGAG 537
 201 LeuLeuSerSerPheAsnPheProTyrLeuHisCysMetTTPHisIleLeuLeuCysLeu 220
 538 CTGCTCTCCCTCTTCTTCACTTCCCTACTGCTGCTGCTGGTGGATATTTCTATCTGCTT 597
 221 AlaAlaTyrLeuGlyCysValCys 228
 598 GCTTCGTACTGGGGCTGTGTGTC 621

RESULT 12

BB660847

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB660847 633 bp mRNA linear EST 26-OCT-2001
 BB660847 RIKEN full-length enriched, 0 day neonate kidney Mus
 musculus cDNA clone D630008P07 5', mRNA sequence.

BB660847.1 GI:16494626

EST.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 633)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,

Konno, H., Koyda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, R., Toyota, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

e mouse tissues.

Location/Qualifiers

1. .633

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="D630008P07"

/tissue_type="kidney"

FEATURES

source

180 eSerGlyLeuTTPTrpThrLeuAla 188

609 TTTGGCTTGTGGTGGACTCTGGCT 633

RESULT 13

BU060795

/dev stage="0 day neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 kidney"
 /notes="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGCGCGCCCAACTCGAGTGTTCCTTTTTCCTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGAGATTCTCGAGTTAATTAATAATCCCTCCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda PLC I."

ORIGIN

Alignment Scores:
 Pred. No.: 2.05e-87 Length: 633
 Score: 916.00 Matches: 169
 Percent Similarity: 93.65% Conservative: 8
 Best Local Similarity: 89.42% Mismatches: 11
 Query Match: 60.94% Indels: 1
 DB: 10 Gaps: 0

US-10-017-410-4 (1-275) x BB660847 (1-633)

QY 1 MetGlyAlaProHisTTPTrpAspGlnLeuGlnAlaGlySerSerGluValAspTTPCys 20
 Db 69 ATGGGCGCCCGCAGCTGCTGGGACCCCTCGGGCTGGCAGTTTCGAGGTGGATGTGTC 128
 QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAnthrIleSerAsnVal 40
 Db 129 GAGGACAACTACACTATCGTGCCTGCCATTGCCGAGTCTTACACACGATCAGCAACGTC 188
 QY 41 LeuPhePheIleLeuProProlIleCysMetCysLeuPheAspGluTyrAlaThrCysLeu 60
 Db 189 TTGTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCGACGATGCAACGTCGCTTC 248
 QY 61 AsnSerAspIleTyrIleLeuTTPTrpLeuValValGlyIleGlySerValTyr 80
 Db 249 AACAGCGGCATCTACTTAATATGACGCTCTCTAGTTGTAGTGGGATGGATCTCTAC 308
 QY 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTTPVal 100
 Db 309 TTCATGCAACGCTGAGTTTCTCTGGGTGAGATGCTTGCATGCACTTGCCATTTCTGGGTT 368
 QY 101 LeuMetCysAlaLeuAlaMetTTPPheProArgTyrLeuProLysIlePheArgAsn 120
 Db 369 CTGATGTGCTTTGGCCATGTGTTTCCAGAGAGGTATTTACCAAGATCTTTTCGGAAT 428
 QY 121 AspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCysLeuAla 140
 Db 429 GACAGGGGCGAGGTTCAGGCGAGTGGTGTGTCTCTGCTGCAATTACACGCTGCTGGCG 488
 QY 141 PheValLysProAlaIleAsnIleSerLeuMetThrLeuGlyValProCysThrAla 160
 Db 489 TTTATCAAGCCGCCCATCAACAATATTTCTCTGATGATTCTGGGACTTCCATGCACTGCG 548
 QY 161 LeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGly-LeuPh 180
 Db 549 CTGCTTGTGTCAGAGCTGAAAGGTGTGACATGTGCTGTGTGTAACTGGGCCCTTTT 608

LOCUS BJ060795 648 bp mRNA linear EST 29-SEP-2003
 DEFINITION laevis cDNA clone XL066nl6 5', mRNA sequence.
 ACCESSION BJ060795
 VERSION BJ060795.1 GI:17500536
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.
 TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shirzuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp
 The information of this clone is available through the following URL.
 http://xenopus.nibb.ac.jp.
 FEATURES
 source
 1..648
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL066nl6"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud library"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,47e-86 Length: 648
 Score: 906.00 Matches: 161
 Percent Similarity: 92.19% Conservative: 16
 Best Local Similarity: 83.85% Mismatches: 15
 Query Match: 60.28% Indels: 0
 DB: 12 Gaps: 0
 US-10-017-410-4 (1-275) x BJ060795 (1-648)
 QY 84 ThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrpValLeuMetCys 103
 Db 5 ACCCTCAGCTTCTCGGGCAGATGCTGGATGAGTTGGCTATCCTGTGGGTGCTATGTCT 64
 QY 104 AlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsnAspArgGly 123
 Db 65 GCGTTGGCCATGTGTTCCCTAAAGACACTTGCCTAAGGGTATTTCCGAATGACAGGTG 124
 QY 124 ArgPheLysValValSerValLeuSerAlaValThrThrCysLeuAlaPheValLys 143
 Db 125 AGTTTAAAGCAGTGTGTTGGGTGATTTCCGGAGTACCACAGCCCTAGCGTTTATCAAG 184
 QY 144 ProAlaIleAsnAsnIleSerLeuMetThrLeuGlyValProCysThrAlaLeuLeuIle 163
 Db 185 CCTGCCATTAAACGATATCTCTGATGATTCCTGGGATACCTTGTACAGCTTACTCATT 244
 QY 164 AlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPheSerGlyLeu 183
 Db 245 GCTGAGCTCAAGAGATGTGACATATATCGGGTGTATTAAGCTTGGAGTGATGCTCGGGTT 304
 QY 184 TrpTrpThrLeuAlaLeuPheCysTrpIleSerAspArgAlaPheCysGluLeuLeuSer 203
 Db 305 TGGTGACGCTGGCGTGTGGCTGTTGGATCAGTGACAAAGCTTTTGTGAAATATATGGTCC 364
 QY 204 SerPheAsnPheProTyrLeuHisCysMetTrpHisIleLeuIleCysLeuAlaIleTyr 223

Db 365 TCTTTCAACTTCCCTACCTCCACTGTGTATGGCACATCTTATCTGCTGCTGCTTAT 424
 QY 224 LeuGlyCysValCysPheAlaTyrPheAspAlaIleAsnGluIleProGluGlnGlyPro 243
 Db 425 CTCGGCTGTGTCTGCTTACTTTCATGCGCATCTGAAATCCAGAACAGGACCT 484
 QY 244 ValIleLysPheTrpProAsnGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
 Db 485 GCCATCAAAATTTGGCTAGTGAAGAGTGGGCTTCATGAGTACCTTATGCTACTTTA 544
 QY 264 LeuCysAlaAsnLysSerSerValLysIleThr 275
 Db 545 CTCCTGCAACACAGAAATCGCCATTAAATAATACA 580
 RESULT 14
 BU220481 847 bp mRNA linear EST 25-NOV-2002
 LOCUS CSEQCHN04 Gallus gallus cDNA clone CHEST4903 5', mRNA
 DEFINITION sequence.
 ACCESSION BU220481
 VERSION BU220481.1 GI:25406614
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 847)
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 FEATURES
 Location/Qualifiers
 1..847
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST4903"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN04"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+);
 Site 1: EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research
 (1996) 791, except that a significantly longer
 reannealing hybridization was used."
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.14e-85 Length: 847
 Score: 896.00 Matches: 157

Percent Similarity: 95.53% Conservative: 14
 Best Local Similarity: 87.71% Mismatches: 8
 Query Match: 59.61% Indels: 0
 DB: 13 Gaps: 0

US-10-017-410-4 (1-275) x BU220481 (1-847)

Qy 97 ValLeuTrrpValLeuMetCysAlaLeuAlaMetTrrpPheProArgArgTrrpLeuProLys 116
 Db 2 ATTCTCTGGTCTTATGTGTGCTCTTCCATGTGTCTTCCCTAGGAGATATCTACCAAGA 61
 Qy 117 IlePheArgAsnAspArgGlyArgPheLysValValValSerValLeuSerAlaValThr 136
 Db :: 62
 Db 62 GTTTTTCGAAATGACAGAGCGGTTAAAGCTGCTGTGGTGTCTCTGGAGTTACT 121
 Qy 137 ThrCysLeuAlaPheValLysProAlaIleAsnAniLeSerLeuMetThrLeuGlyVal 156
 Db :: 122
 Db 122 ACCTGCCTTGGCTTCATTAACCTGCCATCAACATCTCATTANTGACTCTGGGTGT 181
 Qy 157 ProCysThrAlaLeuLeuAlaGluLeuLysArgCysAspAsnMetArgValPheLys 176
 Db 182 CTTTGCACAGCTTACTCATTTGAGTTGAAGAGGTGTGAACACCTCGGTGTGTACAAAG 241
 Qy 177 LeuGlyLeuPheSerGlyLeuTrrpThrLeuAlaLeuPheCysTrrpIleSerAspArg 196
 Db :: 242
 Db 242 CTTGGTCTGTTTTCAGGTCCTTGGTGGAGTGTAGCAGCTTTTCTGCTGGATCAGTGACAAA 301
 Qy 197 AlaPheCysGluLeuLeuSerSerPheAsnPheProTrrpLeuHisCysMetTrrpHisIle 216
 Db 302 CTTTGTGTGAGTCTGTGTCATCTTAACCTTCCCTATTTGACATGTGTATGGCACATT 361
 Qy 217 LeuLeuCysLeuAlaAlaTrrpLeuGlyCysValCysPheAlaTrrpPheAspAlaAlaSer 236
 Db 362 TTGATTTCCTTGGCTTACTAGGCTGTGCTGTTTCTGCTTACTTCGATGCTGCCTCC 421
 Qy 237 GluileProGluGlnGlyProValIleLysPheTrrpProAsnGluLysTrrpAlaPheIle 256
 Db 422 GAGATCCTGTAGCAGGCGCCGTATAAAGTTCTTGCCCAAGTGAAGATGGGCATTTCATT 481
 Qy 257 GlyValProTrrpValSerLeuLeuCysAlaAniLysLysSerSerValLysIleThr 275
 Db 482 GCGGTTCCCTAGCTCACCTCTCTGTGCACACAGAAATCACCGGTGAAGATACACA 538

RESULT 15
 BX842808
 LOCUS
 DEFINITION
 IMAGE:3402865 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BX842808 937 bp mRNA linear EST 11-DEC-2003
 BX842808 NICHU XGC Emb1 Xenopus laevis cdna clone IMAGp998P028304 ;
 IMAGE:3402865 5', mRNA sequence.
 BX842808
 BX842808.1 GI:39733361
 EST.
 Xenopus laevis (African clawed frog)
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 Heil, O., Neubert, P., Peters, M., Radelof, U., Schneider, D.,
 Schroth, A., Korn, B. and Landgrebe, J.
 Xenopus laevis UniGene Set 1 (RZPDLib No. 988)
 Unpublished (2003)
 Contact: Ina Rolfs

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGp998P028304.
 RZPDLib; I.M.A.G.E. cdna Clone Collection (amp- resistant) (RZPDLib
 No. 998) [http://www.rzpd.de/cgi-](http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=998)
[bin/products/showLib.pl.cgi/response?libNo=998](http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=998) [http://www.rzpd.de/cgi-](http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=998)
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel.: +49 30 32639 101

FEATURES

source

1. 937
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGp998P028304 ; IMAGE:3402865"
 /tissue_type="embryo (stage 10)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHU XGC Emb1"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
 Cloned unidirectionally. Primer: Oligo dt. Average insert
 size 1.55 kb. Constructed by Life Technologies. Note: This
 is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:
 Pred. No.: 1,49e-83 Length: 937
 Score: 882.00 Matches: 163
 Percent Similarity: 89.45% Conservative: 15
 Best Local Similarity: 81.91% Mismatches: 20
 Query Match: 58.68% Indels: 1
 DB: 13 Gaps: 0
 US-10-017-410-4 (1-275) x BX842808 (1-937)

Qy 1 MetGlyAlaProHisTrrpTrrpAspGlnLeuGlnAlaGlySerSerGluValAspTrrpCys 20
 Db 338 ATGGCGCGCTTTAAAGATCTGGGAGCATCTTCAGGCAGGCAGCTCAGAAAGTGCAGCTGTGC 397
 Qy 21 GluAspAsnTrrpThrIleValProAlaIleAlaGluPheTrrpAsnThrIleSerAsnVal 40
 Db 398 GAGGACAAATTACACCATAGTGTCTTCAATCCAGAGTTCTTCAATACATACGATCAGCAACATT 457
 Qy 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheAspGluTrrpAlaThrCysLeu 60
 Db 458 TTATTTTGTGTTTACCGCCCATATGATGTGTCTATTCCTCAGTACGACCACTGCTTC 517
 Qy 61 AsnSerAspIleTrrpLeuIleTrrpThrLeuLeuValValValGlyIleGlySerValTrr 80
 Db 518 AACAGTGGGATATATTAATATGACTCTGCTGCTGTGTGTGCTGATTTGGTCCGTGAT 577
 Qy 81 PheHisPheThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaValLeuTrrpVal 100
 Db 578 TTCCATGCCAGCTCAGTCTTTCTCGGCAAAATGCTGGATGAGTTGGCTATCTCTGTGGGTG 637
 Qy 101 LeuMetCysAlaLeuAlaMetTrrpPheProArgTrrpLeuProLysIlePheArgAsn 120
 Db 638 CTAATGCTGATTTAGCCATGTGTGTTCTTAAAGACATCTGCCAAGGATTTTCGAAT 697
 Qy 121 AspArgGlyArgPheLysValValValSerValLeuSerAlaValThrThrCysLeu-Al 140
 Db 698 GACAGGTGGAGGTTTAAAGCAGAGTTGTTGGGGTGTCTTTCGGGAGTCAACACAGCCCTAGGG 757
 Qy 140 aPheValLysProAlaIleAsnAniLeSerLeuMetThrLeuGlyValProCysThrAl 160
 Db 758 GTTTATCAAGCTGCCATTAACGATATCTCCTGATGACATCTGGGATACCATGATACAGT 817
 Qy 160 aLeuLeuIleAlaGluLeuLysArgCysAspAsnMetArgValPheLysLeuGlyLeuPh 180
 Db 818 CTTTACTCATGCTGAGCTCAAGAGATGTGACACATCGGGGTTTACAAGCTTGGCCTAAT 877
 Qy 180 eSerGlyLeuTrrpThrLeuAlaLeuPheCysTrrpIleSerAspArgAlaPhe 198
 Db 878 GTCTGGATTTGGTGGACATTGGCATTGGCCTGTGGATCATCGGACGACCAAGCTTTT 932

Search completed: September 18, 2004, 06:56:45
 Job time : 2475 secs

This Page Blank (uspto)